



ARROYO CENTER

Assessing, Monitoring, and Evaluating Army Security Cooperation

A Framework for Implementation

Angela O'Mahony, Ilana Blum, Gabriela Armenta, Nicholas Burger,
Joshua Mendelsohn, Michael J. McNerney, Steven W. Popper,
Jefferson P. Marquis, Thomas S. Szayna

For more information on this publication, visit www.rand.org/t/RR2165

Library of Congress Cataloging-in-Publication Data is available for this publication.

ISBN: 978-0-8330-9941-9

Published by the RAND Corporation, Santa Monica, Calif.

© Copyright 2018 RAND Corporation

RAND® is a registered trademark.

Limited Print and Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law. This representation of RAND intellectual property is provided for noncommercial use only. Unauthorized posting of this publication online is prohibited. Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Permission is required from RAND to reproduce, or reuse in another form, any of its research documents for commercial use. For information on reprint and linking permissions, please visit www.rand.org/pubs/permissions.

The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest.

RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

Support RAND

Make a tax-deductible charitable contribution at
www.rand.org/giving/contribute

www.rand.org

Preface

Security cooperation is an important instrument of U.S. foreign policy, and is employed extensively to accomplish a diverse set of objectives, such as building relationships that promote U.S. security interests, developing partners' capabilities for self-defense and multinational operations, and securing access for U.S. forces. However, evaluating the effectiveness of security cooperation to meet these objectives has proven elusive.

To help the Army increase the effectiveness of its security cooperation activities and to more effectively evaluate its future security cooperation activities, the U.S. Army Deputy Chief of Staff, G-3/5/7, asked the Arroyo Center to address two questions. First, when can Army security cooperation have the greatest impact? Second, how should the Army assess, monitor, and evaluate security cooperation? Taken together, the answers to these two questions can help the Army prioritize and evaluate future security cooperation activities.

Part I of this report addresses the first question. It focuses on understanding the factors that are associated with security cooperation effectiveness and examines the extent to which recent Army security cooperation activities reflect those factors. Part II addresses the second question. It presents a framework to help the Army implement an assessment, monitoring, and evaluation (AM&E) process in line with good practice and emerging Department of Defense guidance. The framework can enable accountability—for planners, implementers, and stakeholders—and institutionalize a learning process within the Army.

The study results will assist the Army to better plan and execute security cooperation activities. The primary users of this framework will be Army Service Component Command planners. The Army's focus on deepening assessment, monitoring, and evaluation procedures at the security cooperation activity level can serve as a model for the Department of Defense (DoD) to link its strategic guidelines to the operational planning level. The Army's bottom-up operational approach can be a valuable complement for the Office of the Secretary of Defense's top-down strategic approach. Taken together, these approaches form the basis for a comprehensive approach to security cooperation, encompassing strategic, operational, and tactical considerations.

This report documents research and analysis conducted as part of a project titled *Developing a Framework to Assess Expected Impact of Army Security Cooperation*, sponsored by the United States Army Deputy Chief of Staff, G-3/5/7.

The Project Unique Identification Code for the project that produced this document is RAN167298.

This research was conducted within the RAND Arroyo Center's Strategy, Doctrine, and Resources Program. RAND Arroyo Center, part of the RAND Corporation, is a federally funded research and development center (FFRDC) sponsored by the United States Army.

RAND operates under a “Federal-Wide Assurance” (FWA00003425) and complies with the Code of Federal Regulations for the Protection of Human Subjects Under United States Law (45 CFR 46), also known as “the Common Rule,” as well as with the implementation guidance set forth in DoD Instruction 3216.02. As applicable, this compliance includes reviews and approvals by RAND’s Institutional Review Board (the Human Subjects Protection Committee) and by the U.S. Army. The views of sources utilized in this study are solely their own and do not represent the official policy or position of DoD or the U.S. Government.

RAND’s publications do not necessarily reflect the opinions of its research clients and sponsors.

Contents

Preface	iii
Figures	vii
Tables	ix
Summary	xi
Acknowledgments	xxv
Abbreviations	xxvii

PART ONE

Understanding Army Security Cooperation Effectiveness	1
--	---

CHAPTER ONE

Introduction	3
Partner Engagement Is an Important U.S. Foreign Policy Tool	3
Evaluating Security Cooperation Is Challenging but Necessary	4
Objectives, Approach, and Organization	6

CHAPTER TWO

Lessons from Security Cooperation and International Development Assistance	9
Why Compare Security Cooperation and Development?	9
Considerations Affecting International Assistance	11
Approach	13
Strategic Considerations (Political Importance)	14
Strategic Considerations (Economic Importance)	15
Political Considerations	17
Absorptive Capacity Considerations	20
Cultural Considerations	21
Financial Assistance Considerations	22
Conclusion	24

CHAPTER THREE

Army Security Cooperation Empirics	27
Approach	28
Overview	31
Strategic Considerations	33
Political Considerations	36
Absorptive Capacity Considerations	38

Cultural Considerations.....	40
Financial Aid Considerations.....	41
Conclusion.....	42
 PART TWO	
Implementing an Army AM&E Framework.....	43
 CHAPTER FOUR	
AM&E Developments and Implications for the Army	45
AM&E Developments—and Requirements—Are Accelerating.....	45
The Army’s Independent Approach to Security Cooperation AM&E.....	49
Applying Army AM&E Policy and Doctrine to Security Cooperation Processes	52
Potential Improvements in Army AM&E.....	54
 CHAPTER FIVE	
Good Practices in Development AM&E	57
AM&E Trajectory in International Development	58
The Use of AM&E in Practice.....	60
Conclusion.....	76
 CHAPTER SIX	
Framework for Assessing, Monitoring, and Evaluating Army Security Cooperation	77
ASCC Planning for Security Cooperation.....	77
ASCC Initial Country Assessments.....	79
TOC for Army Security Cooperation Activities	80
Army Security Cooperation M&E.....	87
Conclusion.....	97
 CHAPTER SEVEN	
Pilot Tool for Evaluating Security Cooperation Portfolios.....	99
Security Cooperation Portfolio Tool: Inputs and Outputs.....	100
Sample Portfolio Analysis for Bandaria	104
Conclusion.....	113
 CHAPTER EIGHT	
Conclusion	115
 APPENDIXES	
A. What Is Security Cooperation?.....	117
B. Regression Model Methodology	125
C. Dimensions of Partner Capabilities Assessment.....	129
D. Examples of Three Notional Activity-Level TOCs	133
E. Activity-Level AM&E Decision Tree and Supporting Annexes.....	145
F. Works Included in Literature Review.....	155
G. Further Reading Suggestions on AM&E and Impact Evaluation Practices	161
 References	 163

Figures

S.1.	Potential Army AM&E Improvements	xvii
S.2.	Army Security Cooperation AM&E Framework Components	xviii
S.3.	ASCC Country Assessment Dimensions	xix
S.4.	Causal Logic of a TOC	xx
S.5.	Activity-Level M&E Decision Tree.....	xxi
S.6.	Exemplar Activity-Level M&E Decision Tree	xxii
S.7.	First Comparison of Security Cooperation Portfolios.....	xxiv
3.1.	Interquartile Range of Activity Attendance Rates by GCC	32
3.2.	Average Portfolio Share Percentage for Each Activity Type in Each GCC	33
4.1.	New Developments and Requirements Affecting AM&E.....	46
4.2.	Comparing Army Operations and DoD AM&E Constructs	50
4.3.	Potential Army AM&E Improvements	54
4.4.	AM&E and Prioritization	55
5.1.	Pyramid for AM&E.....	62
5.2.	AM&E Cycle.....	63
6.1.	ASCC Security Cooperation Planning Process	78
6.2.	Causal Logic of an Activity-Level TOC	83
6.3.	Components Involved in a TOC	87
6.4.	Activity-Level M&E Decision Tree.....	93
6.5.	Exemplar Activity-Level M&E Decision Tree	94
7.1.	First Comparison of Security Cooperation Portfolios.....	106
7.2.	Comparison of Costs and Presumed Outcomes of Two Alternative Portfolios	107
7.3.	Second Comparison of Security Cooperation Portfolios.....	108
7.4.	Peeling Trajectory of Alternative Specifications.....	110
7.5.	Third Comparison of Security Cooperation Portfolios.....	112
A.1.	Six Categories of Security Cooperation Activities	121
D.1.	Notional TOC for MEDCAP to Djibouti Village	135
D.2.	Notional TOC for Colombian Participants in MCCC.....	139
D.3.	Notional TOC for Southern Accord 2016 (Malawi)	142
E.1.	AM&E Decision Tree for Three Example Activities	146

Tables

S.1.	Factors Associated with Assistance Provision and Effectiveness.....	xiv
S.2.	Variation in Effectiveness of Security Cooperation Across Key Characteristics.....	xvi
2.1.	Factors Associated with Assistance Provision and Effectiveness.....	12
3.1.	Predicted Activity Attendance, as Strategic Importance Varies.....	35
3.2.	Predicted Difference in Portfolio Share, Given Strategic Importance.....	35
3.3.	Predicted Activity Attendance, as Political Characteristics Vary.....	37
3.4.	Predicted Difference in Portfolio Share, as Political Characteristics Vary.....	38
3.5.	Predicted Activity Attendance, as Absorptive Capacity Characteristics Vary.....	39
3.6.	Predicted Difference in Portfolio Share, as Absorptive Capacity Characteristics Vary.....	39
3.7.	Predicted Difference in Portfolio Share, as Cultural Characteristics Vary.....	41
5.1.	Comparison of AM&E Approaches.....	73
6.1.	Information Sources for Partner Assessments.....	80
6.2.	Ease of Producing a TOC.....	82
6.3.	Decision Tree Classification for Different Bandaria Activities.....	95
7.1.	Exemplar Security Cooperation Activity Types.....	100
7.2.	Units for Representing Activity Type Level of Effort.....	101
7.3.	Scale Effect Assumptions for Assurance and Deterrence Objective.....	103
7.4.	Presumed Costs and Potential Effect of Portfolio.....	104
7.5.	Comparison of Components of Two Alternative Portfolios.....	107
7.6.	Density and Coverage Measures for Numbered Boxes in Figure 7.4.....	110
7.7.	Conditions Associated with Box 39 in Figure 7.4.....	111
7.8.	Comparison of Two Alternative Portfolios with Joint Exercise at Different Scales.....	113
B.1.	Activity Participation.....	126
B.2.	Portfolio Share by Activity Type.....	128
F.1.	Works Included in Literature Review.....	155

Summary

The U.S. Army implements security cooperation activities across the globe in support of core U.S. national security and foreign policy objectives. These objectives include building relationships that promote U.S. security interests, developing partners' capabilities for self-defense and multinational operations, and securing overseas access for U.S. forces. Planning for these activities is often robust. Many events are tracked in several ways by multiple stakeholders. And good-news stories along with challenges and lessons are shared at conferences, through emails and memos, and in congressional testimony.

Evaluating the effectiveness of security cooperation in a systematic and rigorous way, however, is challenging, particularly at the Army Service Component Command (ASCC) and Headquarters, Department of the Army (HQDA) levels. First, while security cooperation activities tend to be short term and specific, the strategic theater objectives that security cooperation activities are designed to accomplish are often long term and not well defined. Second, the causal relationships that link security cooperation activities to desired outcomes are often not well understood. Finally, collecting information relevant for evaluating security cooperation effectiveness can be difficult. Although all Army security cooperation activities should be assessed through objective and rigorous after-action reporting and all security cooperation tasks should be evaluated through measures of performance and measures of effectiveness, current Army security cooperation is constrained by poor collection of relevant information.

Regardless of the difficulties inherent in evaluating security cooperation activities, the Army and, more broadly, the Department of Defense (DoD) are facing congressional pressure for increased accountability. In response, the DoD is developing an assessment, monitoring, and evaluation (AM&E) policy for its security cooperation enterprise. In the DoD's AM&E construct, *assessment* refers to baseline assessments for leaders and planners that address questions like these: What is the security environment in which these activities will take place? What are the partner's existing capabilities, and what does the United States want to improve? How well aligned are the partner's interests and values with those of the United States? *Monitoring* means tracking efforts to determine whether inputs (e.g., money and effort) are translating into outputs (e.g., equipment, training, education, and information). These outputs then serve as the basis for tracking progress toward objectives (i.e., outcomes). *Evaluation* examines outcomes and is crucial to understanding what is working and what is not. Success is not ultimately measured by the provision of equipment or training; it is measured by the extent to which security cooperation activities help achieve U.S. objectives. Investments require following up to make sure that they yield the full potential benefits

that were expected.¹ This study provides a framework and a tool to help the Army improve its security cooperation AM&E capabilities in line with emerging DoD guidance.

Study Objectives and Approach

To help the Army strengthen the impact of its security cooperation activities and to more effectively evaluate and prioritize them, the U.S. Army Deputy Chief of Staff, G-3/5/7, asked the Arroyo Center to address two questions. First, when can Army security cooperation have the greatest impact? Second, how should the Army assess, monitor, and evaluate security cooperation?

Part I of this report addresses the first question. We focus on understanding the factors that are associated with security cooperation effectiveness and examine the extent to which recent Army security cooperation activities reflect those factors. We did so in two steps. First, we conducted a review of both security cooperation and international development assistance studies to identify factors in each literature that correspond to when assistance was provided and where assistance was effective. Second, we conducted a statistical analysis of Army security cooperation activities with over 150 countries between 2009 and 2014 to determine the extent to which these activities aligned with the lessons learned from our literature review.

Part II of this report addresses the second question. In it, we present a framework to help the Army implement an AM&E process in line with good practice and emerging DoD guidance. Our objective was to develop a practical framework that will help Army planners determine the effectiveness of security cooperation activities and that is complementary with emerging DoD guidance *and* current Army doctrine. We used three building blocks for this framework: (1) AM&E good practices, based on our review of practices adopted by government agencies and multinational organizations conducting international development assistance; (2) emerging DoD guidance and evolving processes for security cooperation AM&E; and (3) current Army doctrine and procedures for security cooperation AM&E. This framework consists of modular components designed as generalizable templates that are easy for planners to modify according to context or need. The framework can enable accountability—for planners, implementers, headquarters staff, and other stakeholders—and institutionalize a learning process within the Army for the entirety of a project or program's life cycle.

Although these two questions are distinct, they are related. For Army planners to undertake effective AM&E of security cooperation activities, they must have a good causal understanding of the factors that are associated with security cooperation effectiveness. While much has been learned about designing effective international assistance activities, as Part I will make clear, there is still much to be learned. The AM&E framework developed in Part II provides templates with which to further the knowledge learned in Part I.

This report is designed to assist the Army to better plan and execute security cooperation activities. The primary users of this framework will be Army Service Component Command planners. More broadly, the Army's focus on deepening AM&E procedures at the security cooperation activity level can serve as a model for DoD to link its strategic guidelines to the operational planning level. The Army's bottom-up operational approach can be a valuable

¹ Michael J. McNerney, Jefferson P. Marquis, S. Rebecca Zimmerman, and Ariel Klein, *SMART Security Cooperation Objectives: Improving DoD Planning and Guidance*, Santa Monica, Calif.: RAND Corporation, RR-1430-A, 2016.

complement for the Office of the Secretary of Defense's (OSD's) top-down strategic approach. Taken together, these approaches form the basis for a comprehensive approach to security cooperation, encompassing strategic, operational, and tactical considerations.

When Can Army Security Cooperation Have the Greatest Impact?

What the Literature Says

Although important differences exist between security cooperation and international development assistance (e.g., stakeholders and objectives often differ), fundamental similarities exist (e.g., importance of matching assistance to partners' absorptive capacity and needs, leveraging small activities to accomplish strategic objectives). Many of the same types of activities are used in both fields, such as training and material support. Moreover, there is a much deeper literature evaluating the impact of international development assistance than exists for security cooperation. To leverage these similarities and the depth of understanding that exists on development assistance effectiveness, we undertook a literature review across both fields to identify factors that correspond to when assistance was provided and where assistance was effective.

In the literature we examined, six partner nation characteristics and two assistance characteristics emerged as most consistently important for explaining assistance provision and effectiveness. The key partner characteristics are countries' strategic importance (both political importance, captured through their political alignment with their donor, and economic importance, captured through their gross domestic product [GDP] per capita), democratic institutions, domestic stability, absorptive capacity, and cultural similarity.² The two assistance characteristics are quantity of aid and consistency of aid. Table S.1 presents a summary of the relationships between each of these eight factors and assistance provision and effectiveness. As the results presented in Table S.1 demonstrate, although the characteristics associated with greater assistance provision often are similar to those associated with assistance effectiveness, they are not identical. Moreover, previous research has built a deeper understanding of the conditions under which assistance is likely to be provided than when assistance is likely to be effective.

Overall, the security cooperation and development assistance literatures were similar in their findings about what types of partners receive assistance, including whether a partner is considered strategically important, the extent to which a partner has democratic institutions prior to receiving aid, which countries had historically received aid from a donor, and, to a lesser extent, shared culture between donor and recipient countries. A key difference emerged in how security cooperation and development assistance thought about partner GDP per capita as an enabling factor, which is discussed in more detail in Chapter Two. The security cooperation and development assistance literatures diverged more strongly on assessments of aid effectiveness. For example, while cultural similarity did not have a notable impact on the effectiveness of development assistance, security cooperation studies found security cooperation was effective between more culturally similar partner countries. There are a couple of potential explanations for these differences. First, analyses of effectiveness in development assistance and security cooperation have focused on different objectives. The development assistance

² For both security cooperation and development assistance, countries' GDP per capita affects the strategic importance of providing a country with assistance. However, the role of countries' GDP per capita reflects very different strategic objectives.

Table S.1
Factors Associated with Assistance Provision and Effectiveness

	Who Receives Assistance?		Is Assistance Effective?	
	Security Cooperation	Development Assistance	Security Cooperation	Development Assistance
<i>Partner characteristics</i>				
Political alignment with United States	■	■	■	■
GDP per capita	■	■	■	■
Democratic institutions	■	■	■	■
Domestic stability	■	■	■	■
Absorptive capacity	■	■	■	■
Cultural similarity	■	■	■	■
<i>Assistance characteristics</i>				
Quantity of aid	■	■	■	■
Consistency of aid	■	■	■	■

SOURCE: RAND analysis.

Note: ■ positive relationship (strong); ■ positive relationship (weak); ■ negative relationship (strong); ■ negative relationship (weak); ■ mixed relationship; ■ no evidence

literature has historically focused on growth in recipient countries.³ In contrast, security cooperation studies favor measures of military effectiveness, that is, whether a partner was better able to defend itself, to support U.S. military operations, or to provide regional or global support to a desired objective after receiving U.S. support. Second, outcome measures have tended to be easier to measure in development assistance than in security cooperation. This has led to a greater willingness on the part of security cooperation scholars to look beyond narrowly quantifiable metrics to assess effectiveness.

While the literature remains mixed on many aspects of the relationship between individual factors and the success of assistance activities, there are a few broad-stroke conclusions that seem well supported by the findings in both security cooperation and development assistance studies:

1. Countries that are considered strategically important receive more aid than countries that are not.
2. Countries with good governance records receive more aid, though governance is often conflated with stability rather than a measure of democratic principles.
3. Aid is used more effectively in countries with good governance.
4. Countries that previously received aid continue to receive aid.
5. Aid is more effective when it is given consistently over time, but there are diminishing returns on investment.

³ The focus in development studies and development assistance provision has shifted in recent years from a focus on aggregate growth to more intermediate outcomes. We build on lessons from this evolution in Part II to identify best practices for evaluating security cooperation activities.

Taken together, these findings suggest that previous lessons about security cooperation effectiveness are most applicable to understanding outcomes in countries that are strategically important, are enduring partners, and have good governance—countries that are similar to members of the North Atlantic Treaty Organization (NATO). As the United States increases the scope of security cooperation, in terms of both activity and partner types, planners will need to understand not only what has been effective in the past but also how to identify lessons from the broader range of partners the United States is engaging. The framework developed in Part II of this report provides a roadmap for doing so.

How Army Security Cooperation Activities Stack Up

We examined how Army security cooperation in recent years matched the findings we identified based on our literature review of previous security cooperation and development assistance studies. We looked at over 9,000 Army security cooperation activities conducted between 2009 and 2014 and recorded in the DoD-wide Global-Theater Security Cooperation Management Information System (G-TSCMIS). Despite its gaps, G-TSCMIS is DoD's most comprehensive data repository for security cooperation activities.

Our analysis confirmed that countries with certain characteristics are especially likely to participate in security cooperation activities. In particular, the Army was more likely to conduct security cooperation activities with stable democracies with relatively well-funded militaries. However, we consistently found that our results were more complex than anticipated based on our survey of the literature, because different country characteristics corresponded to different kinds of assistance, rather than different volumes of assistance. For example, while country stability is a strong predictor of security cooperation in general, country *instability* is a very strong predictor of humanitarian assistance.

We examined the security cooperation portfolios for over 150 countries to determine how much of each portfolio contributed to addressing different security cooperation objectives to understand the extent to which Army security cooperation is tailored by country and objective. We examined the share of each country's engagement portfolio devoted to each of six broad goals of security cooperation:

1. Outreach—demonstrating goodwill and laying the groundwork for future relations. Examples: humanitarian assistance and transnational threat cooperation.
2. Professionalization—professionalization of local military and stabilization of the local security environment. Examples: human rights training, professional military education, and peacekeeping operations.
3. Operational effectiveness—improving the operational effectiveness of partner militaries. Examples: foreign military financing (FMF), foreign military sales (FMS), and advanced military education.
4. Shared values and priorities—facilitating convergence on mutually desired ends of security cooperation and acceptable ways of accomplishing those ends. Examples: political-military dialogue, senior leader engagement, efforts to build doctrinal consensus, and educational programs to increase support for norms such as democracy and civilian control of the military.
5. Joint operations—improving the ability of U.S. and partner militaries to work closely to accomplish common ends. Examples: interoperability, personnel exchanges, information-sharing, and research and development.

6. Deterrence and reassurance—discouraging potential adversaries from aggressive action, and improving the ability of the United States to win wars if deterrence fails. Examples: combat exercises, access agreements, surveillance of potential aggressors, and logistics.

Table S.2 shows how the amount of security cooperation activities within each country's portfolio varies across strategic objectives depending on partner countries' characteristics and the amount and consistency of aid that the United States provides. A few findings stand out. We found that activities in support of outreach, shared values and priorities, and joint operations were most likely to be held with countries that were dissimilar to the United States. These activities ranged from building nascent relationships to increasing interoperability to combat specific national security challenges. In contrast, the Army undertook deterrence-focused activities with its strongest partners and allies.

Overall, we found that Army security cooperation generally favors three types of countries—those in need of greater engagement (e.g., countries with high domestic instability), those with which the United States would like to improve relations (e.g., countries with low political alignment with the United States), and those with which greater engagement will be most productive (e.g., stable democracies with strong military capability). Diving further into these general patterns, we found that different kinds of engagement are administered to different kinds of countries, based on perceptions of the problem to be solved, and beliefs about what can be accomplished through specific types of engagement. Army security cooperation activities have aligned fairly well with what previous analyses have found contribute to effectiveness in both security cooperation and international development assistance. However, the lack of systematic AM&E across Army security cooperation activities makes it impossible to examine whether, as a whole, previous activities met their objectives effectively.

Table S.2
Variation in Effectiveness of Security Cooperation Across Key Characteristics

	Outreach	Professional- ization	Operational Effectiveness	Values and Priorities	Joint Operations	Deterrence and Reassurance
Political alignment with United States	■	□	□	■	■	■
GDP per capita	■	■	■	■	■	■
Democratic institutions	■	■	■	□	■	□
Domestic instability	■	■	■	□	■	□
Absorptive capacity (high percentage of high school graduates)	■	■	■	■	■	■
Absorptive capacity (high military spending)	■	■	□	□	□	■
Cultural similarity	■	□	■	■	■	■
Large amount of aid	■	□	□	□	■	□
Inconsistent aid	□	■	□	□	■	□

SOURCE: RAND analysis.

Note: ■ positive relationship; ■ positive but statistically insignificant relationship; ■ negative relationship; ■ negative but statistically insignificant relationship; □ no relationship

How Should the Army Assess, Monitor, and Evaluate Security Cooperation?

Within the past ten years, several developments laid the foundations for a more strategic and analytically rigorous approach to security cooperation. For example, Presidential Policy Directive 23, new congressional authorities, a Department of Defense Instruction (DoDI), and improved combatant command (CCMD) and service planning have improved the guidance and processes that help shape Army AM&E.

However, as described in Chapter Four and shown in Figure S.1, the Army has more work to do to improve and standardize the processes that would allow for the effective implementation of these frameworks.

Lessons from the Development Community

As our review of the security cooperation and international development assistance literatures found, fundamental similarities exist between the objectives and activities of security cooperation and those of development assistance. Therefore, as a first step to developing a framework to help the Army implement an AM&E process in line with good practice and emerging DoD guidance, we reviewed international development assistance organizations' AM&E processes. Much like security cooperation planners, development organizations are working to set clear achievable metrics, develop assessment frameworks that are calibrated to program needs and constraints, and increase cooperation with partner institutions, especially in host countries. The development community has developed robust, sophisticated methods for monitoring and evaluating its assistance activities. We identified lessons learned from the development community that are useful guidance for the Army:

- All activity planning should incorporate AM&E.
- AM&E should not be “one size fits all.”
- Return on investment in learning should drive AM&E methods.
- High-quality theories of change (TOCs) are valuable prioritizing and evaluating activities.⁴

Figure S.1
Potential Army AM&E Improvements

Iterate	Army needs iterative dialogue informed by experimentation, learning, and accountability
Lead by doing	Army can shape the DoD's AM&E processes through innovative and effective implementation
Focus on outcomes	Outcomes are less quantifiable and less predictable than output, yet more important
Lead with initial country assessments	Army is well positioned to provide assessments of country's capabilities
Improve reporting	Army can work with ASCCs to strengthen evaluation reporting
Prioritize	Army can work with OSD and CCMDs to ensure limited resources for AM&E are applied rationally through vigorous prioritization

RAND RR2165A-S.1

⁴ A TOC is a conceptual tool that maps out how an activity accomplishes the implementer's desired goals in a particular context. This will be discussed in detail in the following section.

- Impact evaluations are valuable for understanding whether and how an activity achieves desired objectives, but, because they are resource intensive, are not appropriate for all activities.
- Qualitative and quantitative evaluation are useful on their own and even more beneficial when used in combination.
- Strong collaboration with partner nations is crucial for effective AM&E.

Framework for Assessing, Monitoring, and Evaluating Army Security Cooperation

To develop a practical framework that can help Army planners determine the effectiveness of security cooperation activities, that is complementary with emerging DoD guidance *and* current Army doctrine, and that builds on lessons learned from the development community, we identified a set of implementation principles. Framework components should be

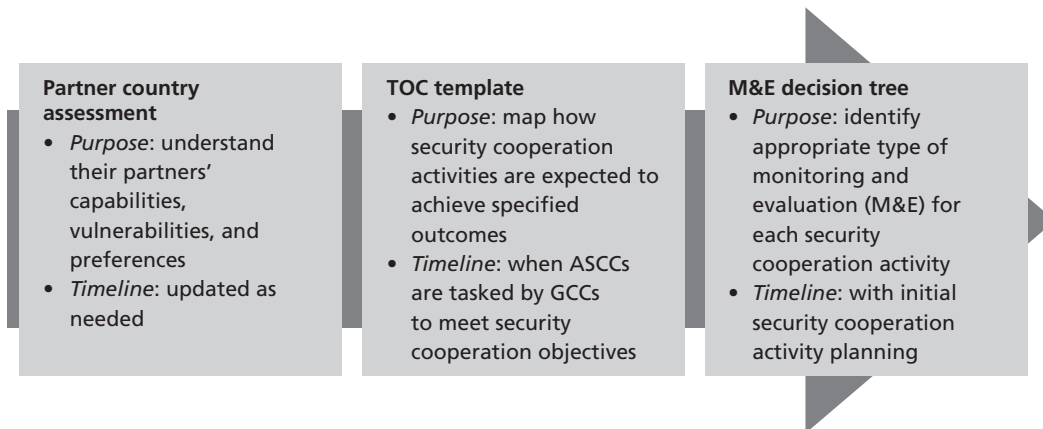
- relevant for Army security cooperation stakeholders, and in particular for ASCC planners, who are the critical link in the Army's security cooperation planning chain
- generalizable across different Geographic Combatant Commands (GCCs), different objectives, and different activities
- consistent with emerging OSD guidance
- designed to enable both accountability and learning.

With these principles in mind, we developed the three framework components shown in Figure S.2.

These components are designed as generalizable templates that are easy to modify according to context or need. These components are meant to serve as a record that can be updated over time, enabling security cooperation planners to plan for and track security cooperation lines of effort. The framework can enable accountability—for planners, implementers, and stakeholders—and institutionalize a learning process within the Army for the entirety of a project or program's life cycle.

As security cooperation continues as an important instrument for the Army to achieve its objectives, the Army needs to know when its security cooperation has been effective. We believe that adopting the Army security cooperation AM&E framework presented in this

Figure S.2
Army Security Cooperation AM&E Framework Components



SOURCE: RAND analysis.

RAND RR2165A-S.2

report will provide a good foundation for deepening the Army's understanding of its security cooperation effectiveness.

ASCC Country Assessment

Country assessments are a key input for ASCC security cooperation planning. To accomplish the GCC objectives, planners need to understand their partners' capabilities, vulnerabilities, and preferences. In turn, the ASCC country assessments inform the GCC strategies and priorities, creating a feedback loop from country assessment to the GCC strategies, through security cooperation tasks. A partner's changing capabilities, characteristics, and strategic context will affect what security cooperation tasks are valuable and feasible. A country assessment also examines how willing and able a partner is to participate in the given activity set and how permissive the environment is for monitoring and evaluation (M&E). Country assessments provide the foundation for activity-specific initial assessments, which are discussed in the following section on TOCs.

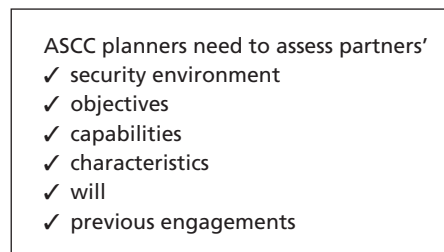
All ASCCs already have procedures in place for conducting country assessments. These assessment procedures have been tailored to meet their region-specific needs within their resource constraints. We expect that ASCCs will not need to substantially alter their country assessment procedures to align with emerging OSD guidance, but recommend that ASCCs examine their current procedures for opportunities to ease any disconnection between the information they are currently collecting and information requests that may emerge from GCCs.

According to emerging OSD guidance, there are several dimensions that need to be included in a partner nation assessment. These are shown in Figure S.3. Although planners should assess all of these dimensions, the depth of knowledge required for an assessment should depend on country priorities—not all partners need to be assessed with the same level of effort. For a non-high-priority country that is stable, relatively minimal assessment may be sufficient. In contrast, high-priority countries or countries that are undergoing periods of flux or uncertainty will require more in-depth assessment. Information sources for lower-priority countries can be primarily open-source documents that are accessible to most ASCC security cooperation planners, while more intensive assessments may require information that is closer held or privileged in some way. For these partners, ASCC planners may want to work with Department of State (DoS) country teams.

TOC for Army Security Cooperation

A TOC is a conceptual tool that maps out how an activity accomplishes the implementer's desired goals in a particular context. It is the basis for verifying whether an organization is

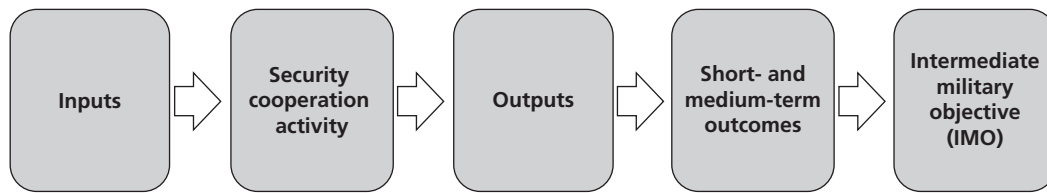
Figure S.3
ASCC Country Assessment Dimensions



SOURCE: RAND analysis.

RAND RR2165A-S.3

Figure S.4
Causal Logic of a TOC



SOURCE: RAND analysis.

RAND RR2165A-S.4

accomplishing expected milestones, for developing lessons learned after activity implementation, and for accountability and transparency purposes. TOCs are helpful for the Army because they ensure planners define and map out the causal link between activities and outcomes or impacts, which in turn supports effective AM&E. Figure S.4 identifies the causal logic that needs to be mapped out in a TOC.

Procedurally, adopting a TOC approach should not require substantial changes to ASCCs' current process. It does, however, require a change in mindset, emphasizing the importance of identifying and critically assessing each link in the causal chain linking security cooperation activities to desired strategic objectives. For a TOC to be an effective tool, planners need to understand the practical implementation issues as well as the conceptual goals. The template and examples provided in this report provide a practical process and illustrations for implementing TOCs for Army security cooperation.

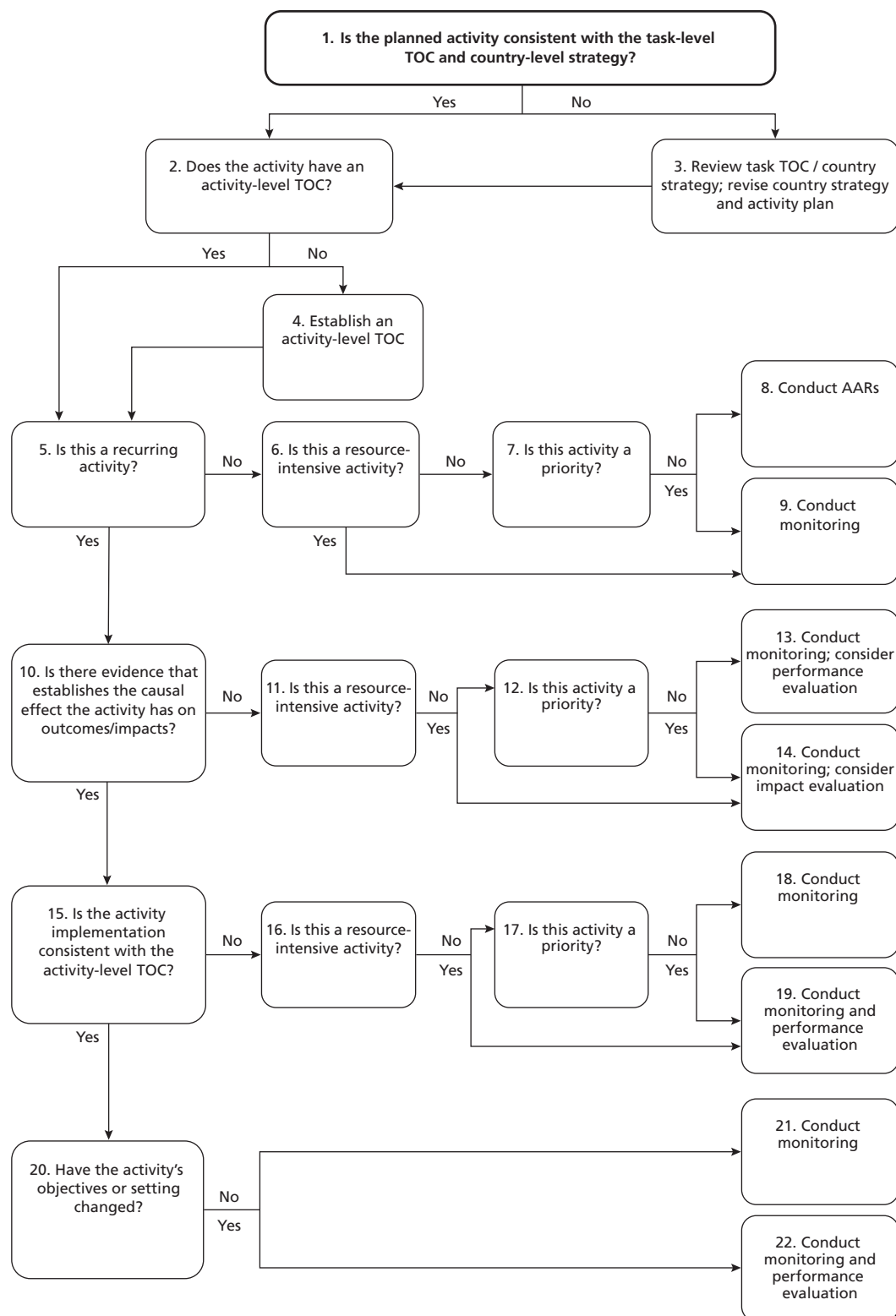
Decision Tree to Guide M&E Decisions

We recommend that all security cooperation activities involve some type of M&E, but the type of M&E used should be determined on a case-by-case basis. Activities that occur frequently, are resource intensive, and for which little is known about the causal relationship between the activity and the desired outcomes require more rigorous or demanding M&E methods. In contrast, security cooperation activities that occur infrequently or are one-offs, those that use few resources, and those for which there is existing evidence of effectiveness need only limited M&E. We consider when planners should adopt four different types of M&E strategies:

- After-action reports (AARs) (only)
- Monitoring
- Performance evaluation
- Impact evaluation.

At the heart of our framework is a decision tree to guide planners' choices on which method to use when considering the appropriate type of M&E for each security cooperation activity. Planners can use this to determine which is the most relevant and appropriate form of M&E: monitoring, performance evaluation, and impact evaluation. *The decision tree is built on the understanding that M&E is always costly, but sometimes it is worth the expense. The decision tree tries to strike the balance between the costs and benefits of M&E.* The decision tree also reminds planners of the fundamentals that need to be in place before implementing any activity-level M&E method. Figure S.5 presents the full decision tree, while Figure S.6 illustrates one pathway leading to a monitoring outcome.

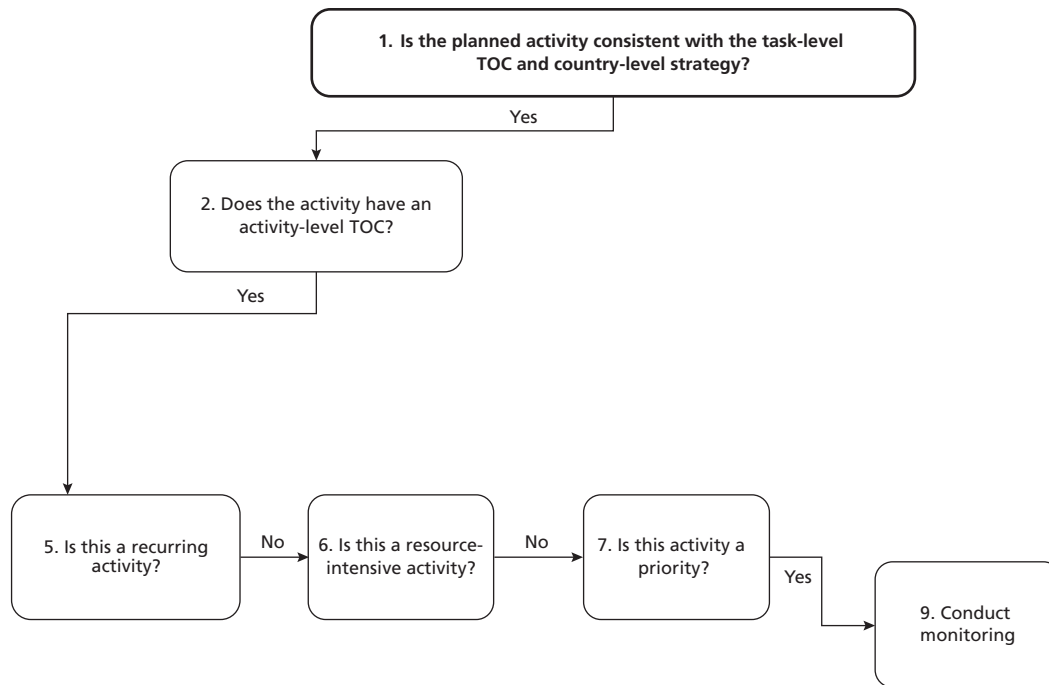
Figure S.5
Activity-Level M&E Decision Tree



SOURCE: RAND analysis.

RAND RR2165A-S.5

Figure S.6
Exemplar Activity-Level M&E Decision Tree



SOURCE: RAND analysis.

RAND RR2165A-S.6

Figure S.6 presents a truncated version of this decision tree. In it, we present a notional M&E strategy for Southern Accord, a multilateral exercise held by U.S. Army Africa. We present this example in more detail in Appendixes D and E. As an initial check, the decision tree ties all M&E decisions to planners' TOCs (boxes 1 and 2). Three decision nodes guided the notional path we outline for Southern Accord. First, is this a recurring activity (box 5)? Southern Accord is a recurring series of multilateral activities, but each one is relatively unique in its execution. As a result, we expect that most planners will view Southern Accord as a nonrecurring activity. In contrast, most training courses would be recurring activities. Second, is this a resource-intensive activity (box 6)? Historically, Southern Accord is a command post and tabletop exercise, which are generally considered medium-cost training exercises—they are much more resource intensive than subject matter exchanges, but are dwarfed by an undertaking such as Pacific Engagement. Third, is this activity a priority? Based on discussions with U.S. Army Africa planners for previous RAND research, Southern Accord was identified as a priority activity.⁵ Southern Accord, and its associated African Land Forces Summit, provides a valuable forum for U.S. Africa and regional leaders to engage. Based on these decision nodes,

⁵ Angela O'Mahony, Thomas S. Szayna, Christopher G. Pernin, Laurinda L. Rohn, Derek Eaton, Elizabeth Bodine-Baron, Joshua Mendelsohn, Osonde A. Osoba, Sherry Oehler, Katharina Ley Best, and Leila Bighash, *Global Landpower Network: Recommendations for Strengthening Army Engagement*, Santa Monica, Calif.: RAND Corporation, RR-1813-A, 2017b.

we recommend that an activity such as Southern Accord incorporate a monitoring strategy in the activity implementation.

Evaluating Security Cooperation Portfolios

The framework for Army security cooperation AM&E presented in this report has focused on security cooperation activities. The Army conducts multiple activities with most partner countries, designed to achieve multiple objectives, which complicates AM&E considerably. As a pilot exercise, in Chapter Six we undertook to explore the possibility that a portfolio-wide view of security cooperation activities could be a useful approach for planners to identify priorities and synergies across security cooperation objectives. We developed a relatively simple proof of concept to examine what a security cooperation planning portfolio analysis tool might entail, and what leverage it could provide planners. The tool was designed in an open-architecture format so that it could provide both transparency and an easily modifiable platform for evaluating outcomes from changing either portfolio inputs or assumptions regarding relationships and therefore outcomes. To serve both these purposes a model was implemented as an Excel Workbook.

We applied the tool to “Bandaria,” the Army’s battle lab, using lines of effort and activities developed for the Army’s Security Cooperation Planners’ Course as the tool’s main inputs and progress toward meeting commonly specified strategic objectives as the tool’s main outputs. For each activity type, users can specify the level of effort to adopt, thereby developing a portfolio of activities to undertake with a partner. We modeled the relationships between security cooperation activities and their impact on achieving strategic objectives in a variety of functional forms that we believe capture the causal relationships in most TOCs for security cooperation. Thus, the tool is designed to enable security cooperation planners and assessors to implement the TOC concept systematically in an operational setting. Rather than acting as a reliable predictor of outcomes, the tool instead supports full exposition of beliefs regarding security cooperation activities and their implications for outcomes. Figure S.7 provides an example of how the tool enables comparison of a notional security cooperation portfolio for Bandaria with a random set of alternative portfolios of comparable cost based on their expected effectiveness meeting six strategic objectives.

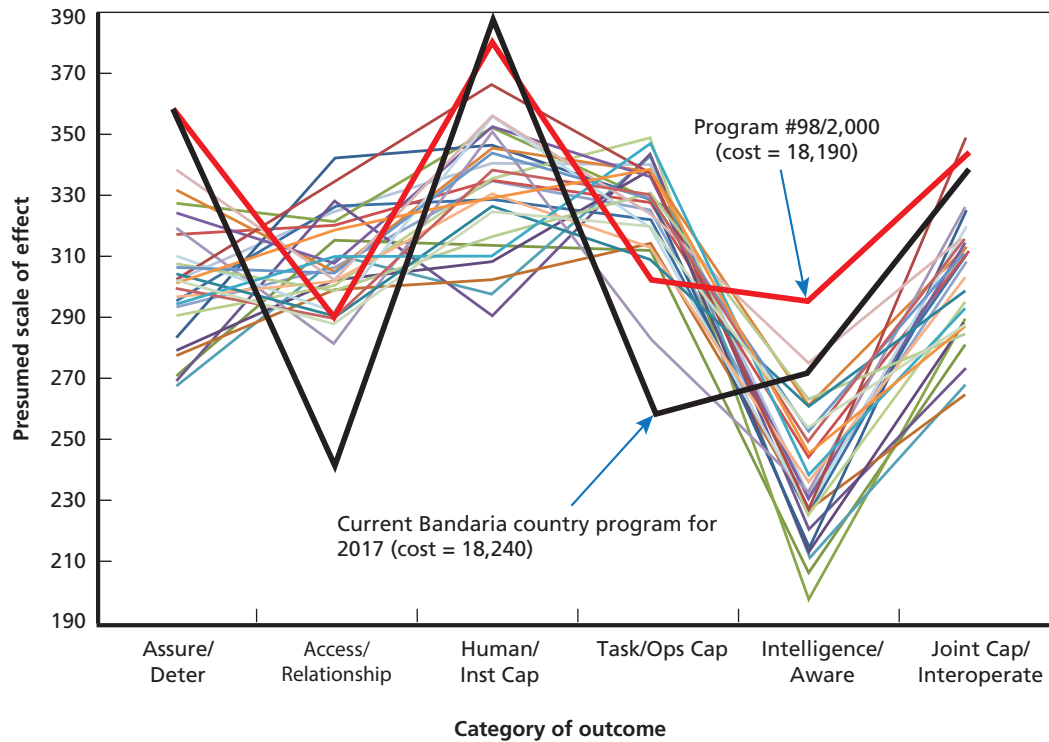
Put another way, the tool allows users to draw detailed inferences about what one would need to believe is likely to be true in order to advocate one course of actions over another in the pursuit of security cooperation objectives. Further, its strength lies not in providing individual “forecasts” but in allowing security cooperation planners and evaluators to set up compound computational experiments and then assisting them in reasoning over the ensembles of results generated by changing assumptions regarding either the appropriate level and mixture of activities within a country portfolio or their presumed effects.

We believe a portfolio analysis tool such as this one could be useful to structure country- and theater-level discussions to align and prioritize security cooperation activities. Explicit consideration of alternative assumptions and their effects on security cooperation outcomes can help identify areas of greatest planning uncertainty.

Conclusion

Leaders in DoD and Congress are directing top-down changes in the planning and execution of security cooperation. By analyzing past activities and established good practices, we have

Figure S.7
First Comparison of Security Cooperation Portfolios



SOURCE: RAND analysis.

NOTE: Presumed effect of current "Bandaria" program compared to 27 random portfolios of LOEs across six outcome measures (all effect assumptions equivalent) with total program cost between 18,000 and 18,240.

RAND RR2165A-S.7

developed in this report a practical, adaptable framework and a portfolio analysis tool to help Army planners and other stakeholders strengthen the impact of security cooperation through improved AM&E.

Acknowledgments

We are grateful to the U.S. Army Deputy Chief of Staff, G-3/5/7, MG William Hix for sponsoring the study. We thank COL Mark Thornhill and COL Joseph Fossey for monitoring the study and providing constructive feedback during its course. We also thank the following people from the office of the U.S. Army Deputy Chief of Staff, G-3/5/7 SSC for working with us throughout our study: LTC Gary Casey, Robert Maginnis, James Freeman, Karen Chipchase, Anthony Lieto, Leslie Thames, and Brenda Wyler. We would also like to thank Jacqueline Bueso-Merriam, Donald Chisholm, Beth Cole, Ann Elizabeth Flanagan, Ana Goicoechea, Alan Gorowitz, Nohemi Lira, Roland Michelitsch, Steve Rader, and Jeffery Tanner for helping us understand good practices undertaken in U.S. government and international development organizations. We would like to thank Solomon Major for sharing data with us. At RAND, we would like to thank Derek Eaton, Bill Gelfeld, Jill Luoto, Rosie Meza, Melinda Moore, Chris Nelson, Lisa Saum-Manning, and Lisa Turner.

Abbreviations

AAR	after-action report
AFRICOM	United States Africa Command
AM&E	assessment, monitoring, and evaluation
ASCC	Army Service Component Command
CCMD	combatant command
CENTCOM	United States Central Command
DoD	Department of Defense
DoDI	Department of Defense Instruction
DoS	Department of State
DOTMLPF-P	doctrine, organization, training, materiel, leadership and education, personnel, facilities, and policy
DSCA	Defense Security Cooperation Agency
EUCOM	United States European Command
FMF	foreign military financing
FMS	foreign military sales
FY	fiscal year
G-TSCMIS	Global-Theater Security Cooperation Management Information System
GCC	Geographic Combatant Command
GDP	gross domestic product
HQDA	Headquarters, Department of the Army
IDB	Inter-American Development Bank
IDD	initial design document
IMO	intermediate military objective
LOA	line of activity

LOE	line of effort
MCC	Millennium Challenge Corporation
MCCC	Maneuver Captains Career Course
M&E	monitoring and evaluation
MEDCAP	Medical Civil Assistance Program
MOE	measure of effectiveness
MOP	measure of performance
NDAA	National Defense Authorization Act
OSD	Office of the Secretary of Defense
PACOM	United States Pacific Command
PMESII	political, military, economic, social, infrastructure, and information
SLE	senior leader engagement
SMART	specific, measurable, achievable, results-oriented, and time-bound
SMEE	subject matter expert exchange
SMS	Strategic Management System
SOUTHCOM	United States Southern Command
TCP	theater campaign plan
TOC	theory of change
UN	United Nations
USAID	United States Agency for International Development
USARPAC	United States Army Pacific
WBG	World Bank Group
WHINSEC	Western Hemisphere Institute for Security Cooperation

PART I

Understanding Army Security Cooperation Effectiveness

Introduction

Partner Engagement Is an Important U.S. Foreign Policy Tool

Engaging with partner countries through security cooperation has become an increasingly important and bipartisan instrument of U.S. foreign policy. The 2014 Quadrennial Defense Report stated that “the United States’ sustained attention and engagement will be important in shaping emerging global trends,” and included “conduct military engagement and security cooperation” as one of the 12 priority missions of the Department of Defense (DoD).¹ The 2015 *National Security Strategy* further deepened these arguments, stating

there are no global problems that can be solved without the United States, and few that can be solved by the United States alone. American leadership remains essential for mobilizing collective action to address global risks and seize strategic opportunities. Our closest partners and allies will remain the cornerstone of our international engagement. Yet, we will continuously expand the scope of cooperation to encompass other state partners, non-state and private actors, and international institutions—particularly the United Nations (UN), international financial institutions, and key regional organizations. These partnerships can deliver essential capacity to share the burdens of maintaining global security and prosperity and to uphold the norms that govern responsible international behavior.²

In his 2017 congressional testimony as nominee for Secretary of Defense, James Mattis said,

Security cooperation extends beyond the military domain. However, the role of the Department of Defense in providing security assistance should be focused on ways to improve the military capacity of other states in order to help them become more reliable and effective partners with the U.S. on security matters. As we do so, we must be prepared to work with even imperfect allies and partner nations to defend our common interests.³

The 2016 and 2017 National Defense Authorization Acts highlight growing congressional interest in security cooperation, as well.⁴

¹ Department of Defense, *Quadrennial Defense Review 2014*, Washington, D.C., 2014, pp. 6, 61.

² The White House, *National Security Strategy*, February 2015, p. 3.

³ James Mattis, “Advance Policy Questions for James N. Mattis, Nominee to Be Secretary of Defense,” 2017.

⁴ See especially U.S. Congress, FY2016 National Defense Authorization Act Section 1202 and FY2017 National Defense Authorization Act Sections 1204 and 1205.

The Army has increased its focus on partner engagement in recent years through its emphasis in the *Army Operating Concept* on engaging regionally,⁵ codification of the engagement warfighting function,⁶ and development of its Global Landpower Network (GLN) concept.⁷ Partner engagement is directly encompassed in two of the Army's current warfighting challenges: shaping the security environment and providing security force assistance.⁸ Security cooperation is a critical instrument for increasing Army engagement with partner countries. The Army undertakes security cooperation activities as an institution and under the aegis of the combatant commanders.

Security cooperation encompasses a diverse set of partner engagement activities, ranging from large-scale combined military exercises to small-scale exchanges of experts or consultations among high-ranking officers. Doctrinally, the Army engages in security cooperation activities to build relationships that promote U.S. security interests, develop partners' capabilities for self-defense and multinational operations, and secure access for U.S. forces.⁹ The growing importance of security cooperation as an instrument of foreign policy reflects the many benefits its proponents have identified. Recent RAND research identified eight benefits that have been ascribed to partner engagement and form the value propositions for the Army's nascent GLN concept.¹⁰ These propositions reflect current Army doctrine, recent Army operational experience, and Army leader statements. Engagement can help

- transition weak partner relationships to longer, enduring, and fruitful ones
- generate knowledge about partners, such as their problems, capabilities, and biases
- address future problems early before they become global
- facilitate local solutions to global problems
- increase operational access to and knowledge of areas and lands
- build interoperability for more robust coalition operations
- increase U.S. capacity and capabilities
- show U.S. strategic intent with clarity and availability to escalate as necessary.

Evaluating Security Cooperation Is Challenging but Necessary

Although security cooperation has become an important instrument of U.S. foreign policy, as the Congressional Research Service has noted, assumptions about security cooperation efficacy are "relatively untested."¹¹ While studies have found support for some security cooperation in some

⁵ TRADOC, *The U.S. Army Operating Concept: Win in a Complex World, 2020–2040*, TRADOC Pamphlet 525-3-1, October 2014b.

⁶ TRADOC, *U.S. Army Functional Concept for Engagement*, TRADOC Pamphlet 525-8-5, February 2014a.

⁷ The GLN concept is presented in "CSA Whitepaper on POTUS Strategy and Global Land Power Network: Delivering National Power on Land Through the GLN," July 23, 2014, and included in TRADOC, October 2014.

⁸ Army Capabilities Integration Center, *Army Warfighting Challenges*, February 27, 2017.

⁹ Headquarters, Department of the Army, *FM 3-22, Army Support to Security Cooperation*, Washington, D.C.: Department of the Army, January 2013.

¹⁰ O'Mahony et al., 2017b.

¹¹ Kathleen J. McInnis and Nathan J. Lucas, *What Is "Building Partner Capacity"? Issues for Congress*, Washington, D.C.: Congressional Research Service, 2015, p. 54.

circumstances,¹² or for some types of security cooperation,¹³ evaluating the effectiveness of security cooperation as part of the Army's and DoD's day-to-day processes has proven elusive.

Security cooperation is challenging to evaluate for three key reasons. First, the strategic theater objectives that security cooperation activities are designed to accomplish are often long term and not well defined. Individual security cooperation activities are rarely able to achieve theater strategic objectives on their own. Multiple activities executed over time are required for security cooperation activities to make progress toward an objective. This spatial and scoping disconnect is exacerbated when the intermediate military objectives guiding Army security cooperation are not "SMART" (specific, measurable, achievable, results-oriented, and time-bound).¹⁴ The Office of the Secretary of Defense (OSD) is currently developing guidance to increase the use of SMART objectives by Geographic Combatant Commands (GCCs), which should increase the ability of Army Service Component Commands (ASCCs) to evaluate their security cooperation activities.¹⁵

Second, the causal relationships that link security cooperation activities to desired outcomes are often not well understood. It is difficult to match security cooperation activities to outcomes. As a result, security cooperation planners and assessors are forced to rely on their theories of how security cooperation ought to affect outcomes in order to plan activities and evaluate whether the activities contributed to desired outcomes. Such theories are often not elaborated, lack empirically supported underpinning, and can conflict with other theories used to plan security cooperation activities.¹⁶

Finally, collecting information relevant for evaluating security cooperation effectiveness can be difficult. Although all Army security cooperation activities should be assessed through after-action reporting and all security cooperation tasks should be assessed through measures of performance (MOPs) and measures of effectiveness (MOEs), current Army security cooperation is constrained by poor collection of relevant information. First, after-action reports (AARs) are often missing, not objective and rigorous, or not accessible by Army security cooperation planners. As the Center for Army Lessons Learned reports,

security cooperation events sometimes are executed with little or no feedback to the ASCC staff or members of the security cooperation office in the U.S. embassy. After action reports and trip reports are vital to the ASCC's strategy and plan development efforts. These AARs and lessons learned, if done well, can inform campaign plans and facilitate ASCC-recommended adjustments to the integrated priority list and comprehensive joint assessment, while refining resource requirements.¹⁷

¹² Christopher Paul, Colin P. Clarke, Beth Grill, Stephanie Young, Jennifer D. P. Moroney, Joe Hogler, and Christine Leah, *What Works Best When Building Partner Capacity and Under What Circumstances?* Santa Monica, Calif.: RAND Corporation, MG-1253/1-OSD, 2013a.

¹³ Michael J. McNerney, Angela O'Mahony, Thomas S. Szayna, Derek Eaton, Caroline Baxter, Colin P. Clarke, Emma Cutrufello, Michael McGee, Heather Peterson, Leslie Adrienne Payne, and Calin Trenkov-Wermuth, *Assessing Security Cooperation as a Preventive Tool*, Santa Monica, Calif.: RAND Corporation, RR-350-A, 2014.

¹⁴ McNerney et al., 2016.

¹⁵ Office of the Secretary of Defense, "Department of Defense Instruction: Assessment, Monitoring, and Evaluation Policy for the Security Cooperation Enterprise," DoDI 5200.01, 2016.

¹⁶ Stephen Watts, Christopher M. Schnaubelt, Sean Mann, Angela O'Mahony, and Michael Schwille, *Pacific Engagement: Forging Tighter Connections Between Tactical Security Cooperation Activities and U.S. Strategic Goals in the Asia-Pacific Region*, Santa Monica, Calif.: RAND Corporation, RR-1920-A, 2017.

¹⁷ Center for Army Lessons Learned, *Security Cooperation: Lessons and Best Practices*, Bulletin: 16-09, March 2016.

Second, security cooperation activity evaluation is often limited to measures of performance and effectiveness that are relatively easy to collect, even though they do not provide information that can address the question of security cooperation effectiveness. Assessors often face a twofold challenge: quantitative data, such as counts of events, may be easy to collect but are not particularly informative; qualitative data, such as an executor's impressions of a partner country's will, may be more informative, but may not be seen as valid.

Regardless of the difficulties inherent in evaluating security cooperation activities, the Army and, more broadly, DoD are facing congressional pressure for increased accountability. The National Defense Authorization Act for Fiscal Year 2017 requires independent assessment of DoD security cooperation programs.¹⁸ In response, DoD is developing an assessment, monitoring, and evaluation (AM&E) policy for its security cooperation enterprise.¹⁹ In DoD's AM&E construct, which we discuss in greater detail in Chapter Four, *assessment* refers to baseline assessments for leaders and planners that address questions like these: What is the security environment in which these activities will take place? What are the partner's existing capabilities, and what does the United States want to improve? How well aligned are the partner's interests and values with those of the United States? *Monitoring* means tracking efforts to determine whether inputs (e.g., money and effort) are translating into outputs (e.g., equipment, training, education, and information). These outputs then serve as the basis for tracking progress toward objectives (i.e., outcomes). *Evaluation* examines outcomes and is crucial to understanding what is working and what is not. Success is not ultimately measured by the provision of equipment or training; it is measured by the extent to which security cooperation activities help achieve U.S. objectives. Investments require following up to make sure that they yield the full potential benefits that were expected.²⁰

This study provides recommendations for the Army to improve its security cooperation AM&E capabilities in line with emerging DoD guidance.

Objectives, Approach, and Organization

Two key questions motivate this study. First, when can Army security cooperation have the greatest impact? Second, how should the Army assess, monitor, and evaluate security cooperation? Taken together, the answers to these two questions can help the Army prioritize and evaluate future security cooperation activities.

Part I of this report focuses on understanding the factors that are associated with security cooperation effectiveness and examines the extent to which recent Army security cooperation activities reflect these factors. We did so in two steps. First, we conducted a review of both security cooperation and international development assistance studies to identify factors in each literature that correspond to when assistance was provided and where assistance was effective. This is reported in Chapter Two. Second, we conducted a statistical analysis of Army security cooperation activities with over 150 countries between 2009 and 2014 to determine

¹⁸ U.S. Congress, *National Defense Authorization Act for Fiscal Year 2017*.

¹⁹ Office of the Secretary of Defense, 2016.

²⁰ McNerney et al., 2016.

the extent to which these activities aligned with the lessons learned from our literature review. This is reported in Chapter Three.

Part II of this report presents a framework to help the Army implement an AM&E process in line with good practice and emerging DoD guidance. Our objective was to develop a practical framework that will help Army planners determine the effectiveness of security cooperation activities and that is complementary with emerging DoD guidance *and* current Army doctrine. We used three building blocks for this framework: (1) AM&E good practices, based on our review of AM&E practices adopted by government agencies and multinational organizations conducting international development assistance; (2) emerging DoD guidance for security cooperation AM&E; and (3) current Army doctrine and procedures for security cooperation AM&E. This framework consists of modular components designed as generalizable templates that are easy to modify according to context or need. These components are meant to serve as living documents that can be modified and updated over time, enabling security cooperation planners to plan for and track security cooperation lines of effort (LOEs). The framework can enable accountability—for planners, implementers, and stakeholders—and institutionalize a learning process within the Army for the entirety of a project or program's life cycle.

Part II is organized as follows. Chapter Four discusses recent developments in security cooperation AM&E across DoD. Chapter Five presents good practices in international development assistance organizations' AM&E processes. Chapter Six builds on the goals for Army security cooperation AM&E identified in Chapter Four and the lessons from international development assistance organizations' AM&E processes identified in Chapter Five to develop a framework that is relevant and actionable for ASCC security cooperation planners. While the framework presented in Chapter Six focuses on evaluating individual security cooperation activities, Chapter Seven presents a pilot tool that can help planners evaluate country-level security cooperation portfolios. Chapter Eight concludes.

Although these two questions are distinct, they are related. For Army planners to undertake effective AM&E of security cooperation activities, they must have a good causal understanding of the factors that are associated with security cooperation effectiveness. While much has been learned about designing effective international assistance activities, as Part I will make clear, there is still much to be learned. The AM&E framework developed in Part II provides templates with which to further the knowledge learned in Part I.

Lessons from Security Cooperation and International Development Assistance

Many studies have discussed security cooperation effectiveness, with most emphasizing its contingent nature and difficulty assessing effectiveness empirically. Indeed, RAND has written many of the studies on the topic. However, much of the security cooperation literature relies on qualitative analyses of a small set of cases and/or quantitative analysis of a small set of factors. In contrast, studies of international development assistance have drawn on a much deeper set of empirical analyses. Moreover, given the longer modern history of international development assistance as a topic of study, analysts have been able to examine how relationships between assistance and partner nation outcomes have changed over time. Although important differences exist between security cooperation and international development assistance (e.g., stakeholders and objectives often differ), fundamental similarities exist (e.g., importance of matching assistance to partners' absorptive capacity and needs, leveraging small activities to accomplish strategic objectives). Not inconsequentially, many of the same types of activities are used in both fields, such as training and material support. In this chapter, we present the results of our survey of both security cooperation and international development assistance literatures. We identified factors in each literature that correspond to when assistance was provided and where assistance was effective. When both literatures identified the same relationships, this provided further support for previous security cooperation findings, such as the importance of matching security cooperation to partners' absorptive capacity. When the literatures identified different results, this suggested that either contextual differences between the two types of assistance may play a role in their differential effectiveness, or that more scrutiny is necessary to understand the potential for gaining insights.

In the next section, we discuss why comparing security cooperation and development assistance is valuable. The following section presents our summary observations from our literature review. We then provide an explanation for how we conducted the literature review and a more technical discussion of our analysis.

Why Compare Security Cooperation and Development?

Funding and support for security cooperation programs and activities have expanded significantly in the past two decades, particularly as the United States has placed more emphasis on building stable partners to combat terrorist and extremist organizations. At the same time, and

partially as a result of this substantial increase in funding, there has been increased pressure on implementing agencies to demonstrate how that money has been used, and whether the results justify the costs.

Assessing security cooperation programs and activities to answer those questions has proven challenging for several reasons. These reasons include the following:

- The sensitivity of the data collected
- Program objectives are often multidimensional and therefore difficult to isolate and assess
- Activities are often implemented in unstable environments where access to program sites, data, and stakeholders is limited or constrained
- Impacts may not be fully realized for several years, creating a mismatch between program implementation and assessment windows.

Because of these and other challenges, the current toolset developed to monitor and evaluate security cooperation programs has been able to show broad trends or levels of correlation between variables, which requires less granular, program-specific data available, but not as able to produce rigorous results that establish a causal relationship between security cooperation and desired outcomes and effects.

Security assistance programs and nonmilitary foreign aid face similar challenges. Both must contend with inconsistent metrics between implementing organizations, a lack of unified assessment systems, and a need to translate those programs and objectives to meet the needs of foreign partners. Therefore, both types of organizations are working toward setting clear, achievable metrics, implementing projects while matching them with assessment frameworks, and doing so in cooperation with partner institutions. At the same time, the analytic community has been refining its recommendations on assessment best practices, recognizing the tremendous value in randomized evaluations, and understanding that randomized evaluations may not always be suitable for assessing the effects of foreign aid, hence recommending a holistic assessment approach.

These parallels between security cooperation and foreign aid are valuable, particularly as they can speak to the determinants of the amount of security cooperation or foreign assistance that a country receives, and the factors that affect the impact of foreign aid or security cooperation. Understanding how nonmilitary foreign aid is distributed, how metrics are formed and accounted for throughout a project's life cycle, and how the outcomes are ultimately determined has potential implications for security cooperation's ability to do the same. It is also useful because nonmilitary foreign aid has a more established history of measuring project progress and outcomes, built on a foundation of development economics and academic research.

For these reasons, this project hypothesized that foreign aid literature may be able to supplement—or, indeed, fill gaps in—knowledge of program assessment where security cooperation has thus far not addressed those topics.¹ More interestingly, where the literature on security cooperation and foreign aid agree, it would strengthen the security cooperation find-

¹ For this comparative analysis we focus on the effect of development assistance on economic growth. Many development programs—and much of the development impact evaluation literature—focus on intermediate outcomes, like health status or school attainment, but the literature on cross-country growth determinants is more apt for our comparison.

ings. Conversely, where the literature within the two fields diverged significantly, “a comparative approach with foreign assistance” would provide new insights into the conditions under which security cooperation findings are valid, and new opportunities for further analysis.

Considerations Affecting International Assistance

Across all the literature we examined, five main strategic areas of consideration, encompassing six partner nation characteristics and two assistance characteristics, emerged as most consistently important for explaining assistance provision and effectiveness.² These considerations and their concomitant characteristics are as follows:

1. **Strategic:** How strategically important a country is to the donor to accomplish its foreign policy objectives. For example, for U.S. military objectives, strategic importance captures whether a partner country is identified in U.S. government strategy and planning documents as a key or critical partner willing and able to operate in coalition with or provide strategic access for U.S. forces, and fits into a larger regional strategy (e.g., a member in the Trans Sahara Counter Terrorism Partnership) or a regional issue (e.g., transnational terrorism). For most development assistance donors, strategic importance reflects whether assistance provided to a country can help alleviate poverty. Thus, with regard to economic capacity, strategic importance for security cooperation and for development assistance can be diametrically opposite. We capture strategic importance in our quantitative analysis through two metrics—one focused on political importance and the second on economic importance. The first is countries’ political alignment. The second is countries’ gross domestic product (GDP) per capita.
2. **Political:** Political considerations reflect two key partner characteristics. The first is how democratic are partner countries’ political institutions. This refers to a country having institutions that respect rule of law and enable equal access to institutions for all citizens, as well as civil-military relationships that emphasize civilian control of the military. The second important political consideration is how stable a partner country is domestically. This refers to the internal stability of a country, including the ability of a federal government to provide law and order and enforce its own borders. Stability may not necessarily refer to a country having democratic institutions.
3. **Absorptive capacity:** Countries’ absorptive capacity captures how much institutional capacity a country currently possesses, and how much additional materiel, funding, training, and other assistance a country can absorb without wasting resources or overwhelming institutions.
4. **Cultural:** Partner countries’ cultural similarity between a donor country and a recipient country captures the extent to which the countries are linked through a shared language, a common history, previous colonial ties, or geography.

² We also examined four other factors—spillover effects, time horizons, aid inertia, and good governance—but did not include them in our final analysis because the characteristics were less well established in the literature basis or could be elided into other characteristics.

5. **Financial assistance:** Financial assistance considerations encompass two assistance characteristics: the quantity and consistency of aid the partner country receives. Quantity of aid is a measure of the amount of aid—whether through funding, materiel, or other support—a country receives on a yearly basis. Consistency of aid is a measure of whether a country receives consistent aid over time.

Table 2.1 presents a summary of the relationships between each of these eight factors and assistance provision and effectiveness. The following sections of this chapter provide a deeper discussion of these relationships. Overall, the security cooperation and development assistance literatures were similar in their findings about what types of partners receive assistance, including whether a partner is considered strategically important, the extent to which a partner has democratic institutions prior to receiving aid, which countries had historically received aid from a donor, and, to a lesser extent, shared culture between donor and recipient countries.

A clear point of divergence between the security cooperation literature and the development literature was how a country's GDP per capita influenced how much aid it received. Development-focused studies have found that countries with lower GDP per capita receive more assistance than countries with higher GDP per capita, while security cooperation-focused studies found a preference for providing assistance to countries with higher GDP per capita. This distinction highlights a key philosophical difference between security cooperation and development activities. Security cooperation looks to increase existing capacity, while development assistance seeks poverty alleviation, first and foremost.

While there are few areas of divergence in the first column, several differences appear more strongly in the analysis of aid effectiveness. For example, while cultural similarity did

Table 2.1
Factors Associated with Assistance Provision and Effectiveness

	Who Receives Assistance?		Is Assistance Effective?	
	Security Cooperation	Development Assistance	Security Cooperation	Development Assistance
<i>Partner characteristics</i>				
Political alignment with United States	■	■	■	■
GDP per capita	■	■	■	■
Democratic institutions	■	■	■	■
Domestic stability	■	■	■	■
Absorptive capacity	■	■	■	■
Cultural similarity	■	■	■	■
<i>Assistance characteristics</i>				
Quantity of aid	■	■	■	■
Consistency of aid	■	■	■	■

SOURCE: RAND analysis.

NOTE: ■ positive relationship (strong); ■ positive relationship (weak); ■ negative relationship (strong); ■ negative relationship (weak); ■ mixed relationship; □ no evidence

not have a notable impact on the effectiveness of development assistance, security cooperation studies found security cooperation was effective between more culturally similar partner countries. There are a couple of potential explanations for these differences. First, analyses of effectiveness in development assistance and security cooperation have focused on different objectives. The development assistance literature has primarily focused on growth in recipient countries. In contrast, security cooperation studies favor measures of military effectiveness, that is, whether a partner was better able to defend itself, to support U.S. military operations, or to provide regional or global support to a desired objective after receiving U.S. support. Second, outcome measures have tended to be easier to measure in development assistance than in security cooperation. This has led to a greater willingness on the part of security cooperation scholars to look beyond narrowly quantifiable metrics to assess effectiveness. Finally, on a procedural note, the low degree of factors associated with development assistance effectiveness may reflect the fact that we focused more heavily on capturing relationships associated with security cooperation effectiveness than development assistance effectiveness. As a result, there may be other factors related to development assistance, such as patronage networks or corruption, that were not captured in this analysis.

In the following sections, we provide a more technical discussion of the similarities and differences between security cooperation and development assistance across each of these considerations, focusing on how differences in how different studies approached their analyses may have affected their results.

Approach

Our approach to identifying factors associated with the assistance provision and effectiveness was primarily to seek out a wide array of literature sources across analyses of both development assistance and security cooperation. We identified relevant literature through key word searches conducted on databases such as RAND publications, dissertation repositories, Google Scholar, and academic journals and databases. Our focus here is primarily on broad-based effectiveness, either of security cooperation or development assistance. There are clearly parallels between the “micro” (project level) aid effectiveness literature and the individual security cooperation activities, but in this section we look more at high-level factors that drive foreign support or explain its effectiveness.

For the aid literature, key word searches included but were not limited to “impact of aid on growth,” “impact of development aid on growth,” and “relationship between aid and growth.” Words such as “investment,” “growth,” and “productivity” were treated as synonyms in this search. The words “assistance” and “aid” were also used as synonyms while searching key words.

For security cooperation, key word searches included but were not limited to “security cooperation and effectiveness,” “impact of security cooperation on growth,” and “relationship between security cooperation and growth” as well as phrases like “measuring impact of security cooperation.” Where present, the words “military foreign aid” and “military aid” were treated as synonyms, and “security cooperation” and “security assistance” were treated as closely identified words, though not exactly alike.

From these separate searches, we found over 70 studies across security cooperation and development literature that discussed which countries received assistance, for what purposes,

and how the aid was ultimately used to meet intended objectives. Articles were then sorted into an Excel spreadsheet that tracked the materials according to publication type, publication date, and authors, along with the various mediating factors discussed in the article, the methodology employed in the article (qualitative, quantitative, mixed methods), and more specific methods, as appropriate. Appendix F lists the works included in our analysis by key factor and methodology. Overall, we found that the security cooperation and aid literature were built from very different types of studies; while sources for the development literature offered an extensive recollection of rigorous academic articles, the security cooperation literature relied primarily on individual and comparative case studies, limited program assessments, and studies built from military doctrine or best practices.

Strategic Considerations (Political Importance)

Throughout our literature review, we found that authors often conflated political and economic importance in their analyses of strategic importance. To make this distinction more explicit, we look at both individually.

Security Cooperation Assistance

Security cooperation analyses find that countries designated by the United States as strategically important based on the actions countries are willing and able to undertake, both by themselves and in coalition with the United States; countries' importance regionally; or countries' participation in a transnational strategic plan, such as partnership through the Trans Sahara Counter Terrorism Partnership program, are more likely to receive assistance than countries not designated as such. However, what the United States identifies as strategically important may also be influenced by the ideological goals of a particular U.S. administration, which can override security-driven goals.³ Furthermore, countries designated as strategically important are more likely to have strong support in U.S. political circles, as well as private industry entities with vested interests, further incentivizing continued foreign aid.⁴ There is mixed evidence to support the claim that aid given to countries designated as strategically important is more effective than aid to countries not designated as such. In fact, the opposite may be true: Bapat argues that aid given to countries to fight terrorism (a strategic interest) creates a moral hazard, providing an incentive to states to pacify terrorist groups only to the extent that they keep receiving aid for an undefined period.⁵ Others suggest that the strategic importance plays the most significant role when strategic interests are matched specifically to activities.⁶

³ James Meernik, Eric L. Krueger, and Steven C. Poe, "Testing Models of U.S. Foreign Policy: Foreign Aid During and After the Cold War," *The Journal of Politics*, Vol. 60, No. 1, 1998.

⁴ Lars Berger, "Guns, Butter, and Human Rights: The Congressional Politics of U.S. Aid to Egypt," *American Politics Research*, Vol. 40, No. 4, 2012.

⁵ Navin A. Bapat, "Transnational Terrorism, US Military Aid, and the Incentive to Misrepresent," *Journal of Peace Research*, Vol. 48, No. 3, 2011.

⁶ Christopher Paul, Michael Nixon, Heather Peterson, Beth Grill, and Jessica Yeats, *The RAND Security Cooperation Prioritization and Propensity Matching Tool*, Santa Monica, Calif.: RAND Corporation, TL-112-OSD, 2013b; McInnis and Lucas, 2015; Stephen Biddle, Ryan Baker, and Julia MacDonald, "Small Footprint, Small Payoff: The Military Effectiveness of Security Force Assistance," Working Paper, September 8, 2015.

Development Assistance

Development aid allocation may be driven by donors' strategic interests, but there is little evidence that strategic importance is a predictor of aid effectiveness on growth. Relatively early work by Burnside and Dollar considers the determinants of aid receipts, and the authors find that aid tends to be provided to serve strategic and commercial interests of donor countries. When separating aid into bilateral and multilateral categories, bilateral aid is more influenced by donors' strategic interests, whereas multilateral aid was found to be largely a function of income level, population, and policies (e.g., a good policy environment).⁷ Burnside and Dollar also found a strong positive relationship between bilateral aid and government consumption. In other words, recipients often tend to divert bilateral aid to government consumption, and this aid is not necessarily assigned to countries with favorable policy environments.

In the years since Burnside and Dollar's work, an expansive literature on the effects of foreign assistance has found little systematic evidence that aid bolsters growth, even taking a recipient country's policy environment into account. Rajan and Subramanian conduct a comprehensive analysis and find that aid has little positive or negative effect on growth, and that effect holds when they control for different policy or geographic environments.⁸

In more recent work on aid allocation, Dreher, Sturm, and Vreeland examine the determinants of donor generosity and, consistent with Burnside and Dollar, they find that if a recipient country happens to be a temporary member of the UN Security Council, then the probability of receiving International Monetary Fund program support increases by about 20 percent.⁹ They find similar results for World Bank aid allocations for countries that become part of the UN Security Council.

Strategic Considerations (Economic Importance)

Security Cooperation

Countries with greater GDP per capita are generally perceived as more strategically important than countries with lower GDP per capita. That said, the literature on the empirical relationship between partner nation GDP per capita and security assistance is surprisingly sparse. This may be because partner nation GDP per capita is more commonly combined with other variables, like governance or absorptive capacity, rather than being taken alone as an indication of a nation's value as a military partner. However, the literature makes clear that a nation's relative GDP per capita says very little by itself about its willingness to invest in its military. Relative to their gross national product, there are a number of countries with low GDP per capita that invest disproportionately in their armed forces, and a number of countries with high GDP per capita that underinvest in military development and maintenance. It is also true that the United States tends to work with countries at all income levels. Thus, an MOE

⁷ Craig Burnside and David Dollar, "Aid, Policies and Growth," *American Economic Review*, Vol. 90, No. 4, 2000.

⁸ Raghuram G. Rajan and Arvind Subramanian, "Aid and Growth: What Does the Cross-Country Evidence Really Show?" *The Review of Economics and Statistics*, Vol. 90, No. 4, 2008.

⁹ Axel Dreher, Jan-Egbert Sturm, and James Raymond Vreeland, "Does Membership on the UN Security Council Influence IMF Decisions? Evidence from Panel Data," *CESifo Working Paper Series No. 1808*, August 2006; Axel Dreher, Jan-Egbert Sturm, and James Raymond Vreeland, "Development Aid and International Politics: Does Membership on the UN Security Council Influence World Bank Decisions?" *Journal of Development Economics*, Vol. 88, No. 1, 2009.

ought to focus more on how a partner prioritizes military spending, and whether assistance offered is at the appropriate level for that partner.

There is some evidence to suggest that countries with higher GDP per capita use military assistance more effectively, but this probably highlights the importance of absorptive capacity rather than income. Further, while there is an implicit assumption in some papers that military assistance creates economic growth, this may not be true. DeRouen and Heo argue that, after controlling for technological progress and externalities, military spending actually hinders modernization and has no positive economic benefits.¹⁰

However, higher amounts of assistance tend to correspond to higher effectiveness of the assistance rendered.¹¹ Assistance is also more likely to be effective when countries are more dependent on higher levels.¹² That said, other studies suggest diminishing returns over time on investment at all quantity levels. Thus, designing programs that adapt over time and carefully conducting ongoing project assessments are seen as necessary to avoid assistance stagnation.

The literature strongly suggests that countries that previously received military assistance are significantly more likely to receive future military assistance. There are many good reasons for this: continuity and predictability in funding, particularly in larger programs, may help to ensure desired results over multiyear implementation periods. From a bureaucratic perspective, it is significantly easier to continue funding programs by default at the same levels as the previous year, rather than constantly reassessing allocations. As with quantity of assistance, maintaining continuity in assistance is seen largely as a positive when it comes to assistance effectiveness, particularly when paired with a sustainment and maintenance plan.¹³ However, assistance that is not reviewed and adjusted over time may become misaligned with current needs, and thereby less effective. Some papers also suggested that maintaining consistent and predictable assistance to recipient nations from year to year may diminish the influence the United States has with respect to that country. However, others suggest that initial influence itself is overestimated and assistance continuity is irrelevant.

Development Assistance

In contrast to the relationship between countries' GDP per capita and receiving security cooperation, in which countries with greater GDP per capita are seen as more strategically important partners, for development assistance donors, countries with lower GDP per capita are the strategic priority. This reflects the fact that development assistance seeks poverty alleviation, first and foremost. Overall, studies tend to find that lower-income countries seem to receive more aid than higher-income countries,¹⁴ and aid allocated to countries with lower GDP per

¹⁰ Karl DeRouen and Uk Heo, "Modernization and the Military in Latin America," *British Journal of Political Science*, Vol. 31, No. 3, 2001.

¹¹ John Sislin, "Arms as Influence: The Determinants of Successful Influence," *The Journal of Conflict Resolution*, Vol. 38, No. 4, 1994.

¹² Beth Elise Whitaker, "Compliance Among Weak States: Africa and the Counter-Terrorism Regime," *Review of International Studies*, Vol. 36, 2010.

¹³ Jennifer D. P. Moroney, David E. Thaler, and Joe Hogler, *Review of Security Cooperation Mechanisms Combatant Commands Utilize to Build Partner Capacity*, Santa Monica, Calif.: RAND Corporation, RR-413-OSD, 2013.

¹⁴ Paul Collier and Anke Hoeffler, "Aid, Policy and Growth in Post-Conflict Societies," *European Economic Review*, Vol. 48, No. 5, 2004, pp. 1125–1145; Burnside and Dollar, 2000; World Bank, "Country Policy and Institutional Assessment. Frequently Asked Questions," 2011.

capita is particularly likely to be multilateral rather than bilateral.¹⁵ In addition, there seems to be a negative association between initial income per capita and subsequent growth, but the evidence identified to support this finding is very limited.¹⁶

Limited evidence was identified to support the hypothesis that economic growth is related to initial income differentials, but the evidence identified does point toward a negative association between initial income per capita and subsequent growth.¹⁷ This finding is consistent with the assumption of conditional convergence, in which poorer economies tend to grow faster than richer economies after they received foreign aid, but they all eventually converge in terms of income per capita.

Political Considerations

Security Cooperation Assistance

A number of security cooperation activities undertaken by the United States focus on creating professional, effective military partners. From the U.S. perspective, professionalization of the military requires the institutionalization of values enshrined in the U.S. military: acceptance of civilian authority over the military, respect for the rule of law, and a commitment to human rights, among others. Because of these stated preferences, we might expect to find that the United States prefers to provide assistance to partners that already have similar values and governance systems. The literature reveals a more complex picture.

In the Cold War era, states that were economically and politically aligned with the United States tended to receive bilateral military assistance regardless of that state's human rights record.¹⁸ In the post-Cold War era, U.S. governments have usually placed democratic values above economic need or political alignment when making military assistance decisions.¹⁹ Cingranelli and Pasquarello further assert that demonstrated respect for human rights affects the decision to provide military assistance significantly, but has less effect on how much military assistance to provide.²⁰ However, other scholars reassessed Cingranelli and Pasquarello's method, and concluded there was no evidence that the United States distributes military assistance according to human rights practices of recipient states.²¹

In terms of whether assistance given to countries with governance profiles similar to our own is more effective, the literature is similarly mixed and less complete. It is difficult to make

¹⁵ World Bank, 2011; Burnside and Dollar, 2000.

¹⁶ Barry P. Bosworth and Susan M. Collins, "The Empirics of Growth: An Update," *Brookings Papers on Economic Activity*, Vol. 2, 2003.

¹⁷ Bosworth and Collins, 2003.

¹⁸ Clair Apodaca and Michael Stohl, "United States Human Rights Policy and Foreign Assistance," *International Studies Quarterly*, Vol. 43, No. 1, 1999.

¹⁹ Shannon Lindsey Blanton, "Instruments of Security or Tools of Repression? Arms Imports and Human Rights Conditions in Developing Countries," *Journal of Peace Research*, Vol. 36, No. 2, 1999.

²⁰ David L. Cingranelli and Thomas E. Pasquarello, "Human Rights Practices and the Distribution of U.S. Foreign Aid to Latin American Countries," *American Journal of Political Sciences*, Vol. 29, No. 3, 1985.

²¹ David B. Carleton and Randall W. Stohl, "The Role of Human Rights in U.S. Foreign Assistance Policy: A Critique and Reappraisal," *American Journal of Political Science*, Vol. 31, No. 4, 1987; James M. McCormick and Neil Mitchell, "Is U.S. Aid Really Linked to Human Rights in Latin America?" *American Journal of Political Science*, Vol. 32, No. 1, 1988.

an exact determination about what role governance plays in informing U.S. decisions, in part because the term “governance” can refer to a number of things. In some contexts, governance is a proxy term for “democracy,” and in other contexts it refers to the overall strength of institutions and how willing and able those institutions are to create and enforce laws. Thus, an autocratic regime may have very high governance levels while simultaneously registering very low on democratic principles and practices. Moreover, security cooperation can promote positive values. Atkinson finds that military-to-military contact has a positive, liberalizing effect on other states, which she attributes to a socialization process.²²

Stability may be one area where the research community and the Army’s respective approaches to assessing security cooperation diverge significantly, in terms of both how stability is defined and measured, and whether stability is ultimately beneficial for military assistance. The research community has in fact done very little work on whether the presence of stability (or instability) affects the United States’ decision to provide military assistance. The empirical analysis on support to different types of regimes shows the United States may in fact be more agnostic when it comes to a preference for stable regimes, but because stability can be positive or negative. Just as in the governance example, countries can be very stable, oppressive autocracies or very unstable fledgling democracies, each of which presents different problems and opportunities for U.S. military assistance.

Significantly more research has been done on whether security cooperation activities have a stabilizing effect in partner nations, though the conclusions are very mixed. Jones et al. suggest that the United States may have the greatest impact in countries that are transitioning from authoritarian to a more open system of governance, which is also when countries tend to be at their least stable.²³ Several studies conclude that providing arms and military support to developing countries—particularly those with a large, significantly poor population—is more likely to result in negative effects, including a greater occurrence of coups, political repression, and prolonged military rule.²⁴ Mehler adds that attempting security sector reform before fundamental political and social changes have occurred will have limited effect, particularly in cases with significant internal cleavages in the security sector.²⁵ Creasey, Rahman, and Smith found that nation-building endeavors during conflict can positively affect economic growth, and that there are strong complementarities between economic assistance and military assistance during conflict.²⁶ McNerney et al. also show that security cooperation has a statistically significant relationship with a reduction in state fragility, but emphasize that the extent to which this is true is dependent on the characteristics of recipient state and the type of security

²² Carol Atkinson, “Constructivist Implications of Material Power: Military Engagement and the Socialization of States, 1972–2000,” *International Studies Quarterly*, Vol. 50, No. 3, 2006.

²³ Seth G. Jones, Olga Oliker, Peter Chalk, C. Christine Fair, Rollie Lal, and James Dobbins, *Securing Tyrants or Fostering Reform? U.S. Internal Security Assistance to Repressive and Transitioning Regimes*, Santa Monica, Calif.: RAND Corporation, MG-550-OSI, 2006.

²⁴ Talukder Maniruzzaman, “Arms Transfers, Military Coups, and Military Rule in Developing States,” *The Journal of Conflict Resolution*, Vol. 36, No. 4, 1992; Blanton, 1999.

²⁵ Andreas Mehler, “Why Security Forces Do Not Deliver Security: Evidence from Liberia and the Central African Republic,” *Armed Forces & Society*, Vol. 38, 2012.

²⁶ Ellyn Creasey, Ahmed S. Rahman, and Katherine A. Smith, “Does Nation Building Spur Economic Growth?” *Economic Inquiry*, Vol. 53, No. 1, January 2015.

cooperation assistance provided.²⁷ Ruby and Gibler further show that officer attendance at professional military education courses has an important stabilizing effect on a country once those officers return home, suggesting a kind of institutional socialization process occurs that encourages values like civilian control of the military.²⁸ In keeping with these mixed results, Blair argues that many of the failures of security cooperation to be effective reflect the lack of integration and coordination both within and across security cooperation providers.²⁹

Development Assistance

Good governance positively influences the amount of aid that a country receives and its economic growth. In particular, multilateral aid eligibility was found to be determined by a country's governance capacity. For instance, World Bank loans and grants to the poorest nations in the world (International Development Association countries) are allocated based on country performance ratings, and these ratings capture the quality of the countries' policies and institutional arrangements.³⁰ Another example is foreign aid eligibility by the Millennium Challenge Corporation (MCC). MCC's aid is allocated to countries with good governance policies, or to countries that have been implementing policies fostering economic growth.³¹

In the context of postconflict settings, Collier and Hoeffler developed ordinary least squares and panel data analyses to understand the determinants of aid provision. They found that donors seem to have an inclination to provide aid to countries that recently came out of conflict, and that a good policy environment also increases the amount of aid that these postconflict countries usually receive.

Mixed results were found in the literature that explained the relationship between good governance capacity and economic growth. Some research suggests that good implementation capacity is essential for project success in postconflict settings,³² while other studies suggest that good institutional environments in aid recipient countries have no effect on economic growth.³³ Collier and Hoeffler provide clarity on these findings. These authors suggest that in years one through three following conflict, country institutional capacity is not greater than average, but implementation capacity seems to increase or even double in years four through ten following conflict. Controlling for time horizons seems therefore to be appropriate when analyzing the impact of aid on growth.

Once conflict ends, donors rush to provide aid to countries that just left conflict, but this amount of aid seems to decrease quickly after the first two years, and after the sixth year it does not seem to be significantly different from the amount of aid received by countries

²⁷ McNerney et al., 2014.

²⁸ Tomislav Z. Ruby and Douglas Gibler, "US Professional Military Education and Democratization Abroad," *European Journal of International Relations*, Vol. 16, No. 3, 2010.

²⁹ Dennis Blair, *Military Engagement: Influencing Armed Forces Worldwide to Support Democratic Transitions, Volume 1: Overview and Action Plan*, Washington, D.C.: Brookings Institution Press, 2013.

³⁰ World Bank, 2011.

³¹ Millennium Challenge Corporation, *On the Cutting Edge of Aid Effectiveness. A Primer: Lessons Learned from the Millennium Challenge Corporation*, 2011, p. 4.

³² Marguerite Duponchel, Lisa Chauvet, and Paul Collier, "What Explains Aid Project Success in Post-Conflict Situations?" *World Bank Policy Research Working Paper Series*, 2010.

³³ Rajan and Subramanian, 2008.

in nonconflict settings.³⁴ Collier and Hoeffler hypothesize that the sudden decrease in donor generosity responds to political interests:

Politicians like to respond to the newsworthy event of peace, but post-conflict newsworthiness rapidly fades. Faced with these political disturbances, aid bureaucracies try to get aid allocations back to normal and so rapidly taper out these special allocations once the initial pledges have been met.³⁵

One form of instability is ethnic tensions. Easterly and Levine conducted a regression analysis of a sample of African countries and found that lower ethnic fractionalization (i.e., higher stability) is associated with higher economic growth in African countries.³⁶ They found that one standard deviation in ethnic fractionalization is associated with an increase of 0.4 standard deviations in an African country's income.

Similar results were found by Dutta, Mukherjee, and Roy, who studied a panel of 120 countries from 1979 to 2008, and used fixed effects and generalized method of moments for their analysis.³⁷ They found a positive and significant effect of aid on domestic investments while controlling for a country's political stability, but the returns of aid on growth diminish at higher levels of aid provided for different levels of political stability.

Rajan and Subramanian used instrumental variables and generalized method of moments to study aid effectiveness in stable versus unstable environments. These authors found no conclusive evidence on whether stable environments are facilitators of aid effectiveness. Their results varied from positive to no effect depending on the sample of countries or methods that they used.

Absorptive Capacity Considerations

Security Cooperation

Security cooperation literature emphasizes the importance of correctly matching the level of assistance provided to a country with its ability to absorb—that is, institutionalize and sustain—that assistance. The reasons given for increasing effectiveness through matching absorptive capacity to need include encouraging partner nations to use resources efficiently—or, in other terms, de-incentivizing unnecessary waste in terms of materiel resources or financial resources, while creating institutional learning that ensures resources are used properly and sustainably for the foreseeable future. Measurements of absorptive capacity in partner nations should also take institutional strength into consideration, accounting for partners that are less or more able to partner effectively with the United States.

However, while security cooperation literature takes absorptive capacity as a given, very few attempts have been made to categorize, measure, or assess the level of absorptive capacity a

³⁴ Collier and Hoeffler, 2004.

³⁵ Collier and Hoeffler, 2004, p. 1137.

³⁶ William Easterly and Ross Levine, "Africa's Growth Tragedy: Policies and Ethnic Divisions," *The Quarterly Journal of Economics*, Vol. 112, No. 4, 1997.

³⁷ Nabamita Dutta, Deepraj Mukherjee, and Sanjukta Roy, "Re-Examining the Relationship Between Domestic Investment and Foreign Aid: Does Political Stability Matter?" *International Review of Applied Economics*, Vol. 29, No. 3, 2015.

partner actually possesses.³⁸ It may be the case that absorptive capacity as an independent variable suffers from the same definitional problems as governance, in that absorptive capacity has been used as a catchall phrase to measure all of the following: equipment, organizational characteristics, readiness, the extent of existing training, technological sophistication, education, language abilities, and doctrine.³⁹ What does seem quite evident in the security cooperation literature is that the intuition that absorptive capacity matters is correct.⁴⁰ The challenge is in determining the best way to measure and report the absorptive capacity of a potential partner in order to tailor assistance most effectively.

The desire to properly assess absorptive capacity of partners—and thus determine which partners should receive what kinds of aid programs, and at what level of monetary and materiel commitment—requires greater attention to developing communication channels with partner governments. While the Army has historically primarily relied on its own measures and metrics for assessing absorptive capacity, security cooperation is moving toward emphasizing partner buy-in and ownership. Creating institutional mechanisms for discussing and comparing aspects of absorptive capacity with counterparts in foreign countries will better serve the academic community's conclusion that absorptive capacity matters.

Development Assistance

There is no evidence identified in the aid literature regarding the impact of military absorptive capacity on donor generosity or economic growth when absorptive capacity is narrowly defined through a military lens. However, when the definition is expanded to include factors like a capacity for institutional learning, training, and bureaucratic efficiency, the literature shows that assistance is more effective where absorptive capacity is higher. Feeny and de Silva argued that absorptive capacity is “limited by the following constraints: (i) human and physical capital constraints; (ii) policy and institutional constraints; (iii) macroeconomic constraints; (iv) deficiencies in the manner in which the international donor community delivers its foreign assistance; and (v) social and cultural constraints.”⁴¹ Any number of these constraints can place a limit on the effectiveness of additional aid, after donors have initially provided assistance.

Cultural Considerations

Security Cooperation

Cultural ties between donor and recipient nation—defined as historical ties (colonial or otherwise) and shared language—correlate strongly with the likelihood that a state will receive aid. There are practical reasons for this—former colonies tend to retain parallel institutions due to historical military presence, requiring fewer initial resources from the donor nation to

³⁸ Dafna H. Rand and Stephen Tankel, *Security Cooperation & Assistance: Rethinking the Return on Investment*, Washington, D.C.: Center for a New American Security, 2015.

³⁹ Robert D. Lamb and Kathryn Mixon, *Rethinking Absorptive Capacity: A New Framework Applied to Afghanistan's Police Training*, Washington, D.C.: Center for Strategic & International Studies, 2013.

⁴⁰ Paul et al., 2013b.

⁴¹ Simon Feeny and Ashton de Silva, “Measuring Absorptive Capacity Constraints to Foreign Aid,” *Economic Modeling*, Vol. 29, No. 3, 2012.

be devoted to cultural awareness training.⁴² In combination with shared language, cultural ties make partnering with a country more attractive and less burdensome.⁴³ While the finding is relatively intuitive, its implications are important: surveys and informal studies suggest that the Army's current foreign language and culture capacity is fairly limited. This may be artificially limiting the number of partner nations with which the United States works, as it has limited capacity to engage with partners outside its cultural lineage.⁴⁴ Military assistance would be more effective if there was an emphasis—before program implementation—on understanding the cultural context of all assistance activities, perhaps including mapping out all relevant actors and their possible impact on project success.⁴⁵

Development Assistance

Many studies identify donor-recipient cultural similarity as a factor in the receipt of aid, but no academic papers implicated it in aid effectiveness. In theory, cultural proximity might increase the typical amount of aid because donors might prefer to give to their former colonies, even when those countries are fragile. Development assistance scholars developed a series of statistical models to assess the relationship between cultural similarity and donor generosity. Findings varied throughout the literature. Rajan and Subramanian, and Fuchs, Dreher, and Nunnenkamp identify the determinants of the amount of aid received.⁴⁶ These papers indicate that having colonial ties or shared language with the donor could have a significant impact on donor aid efforts. Aid research papers did not show any evidence of cultural similarity as a predictor of growth.

Financial Assistance Considerations

Security Cooperation

Higher amounts of assistance tend to correspond to higher effectiveness of the assistance rendered.⁴⁷ Assistance is also more likely to be effective when countries are more dependent on higher levels.⁴⁸ That said, other studies suggest diminishing returns over time on investment

⁴² Angela O'Mahony, Thomas S. Szayna, Michael J. McNerney, Derek Eaton, Joel Verneti, Michael Schwille, Stephanie Pezard, Tim Oliver, and Paul S. Steinberg, *Assessing the Value of Regionally Aligned Forces in Army Security Cooperation*, Santa Monica, Calif.: RAND Corporation, RR-1341/1-A, 2017a.

⁴³ Allison Abbe, Lisa M. V. Gulick, and Jeffrey L. Herman, *Cross-Cultural Competence in Army Leaders: A Conceptual and Empirical Foundation*, Arlington, Va.: United States Army Research Institute for the Behavioral and Social Sciences, 2007.

⁴⁴ Allison Abbe, *Building Cultural Capability for Full-Spectrum Operations*, Arlington, Va.: United States Army Research Institute for the Behavioral and Social Sciences, 2008.

⁴⁵ International Security Advisory Board, *Report on Security Capacity Building*, Washington, D.C.: International Security Advisory Board, 2013.

⁴⁶ Rajan and Subramanian, 2008; Andreas Fuchs, Axel Dreher, and Peter Nunnenkamp, "Determinants of Donor Generosity: A Survey of the Aid Budget Literature," *World Development*, Vol. 56, April 2014.

⁴⁷ Sislin, 1994.

⁴⁸ Whitaker, 2010.

at all quantity levels. Thus, designing programs that adapt over time and carefully conducting ongoing project assessments are seen as necessary to avoid assistance stagnation.

The literature strongly suggests that countries that previously received military assistance are significantly more likely to receive future military assistance. There are many good reasons for this: continuity and predictability in funding, particularly in larger programs, may help to ensure desired results over multiyear implementation periods. From a bureaucratic perspective, it is significantly easier to continue funding programs by default at the same levels as the previous year, rather than constantly reassessing allocations.

As with quantity of assistance, maintaining continuity in assistance is seen largely as a positive when it comes to assistance effectiveness, particularly when paired with a sustainment and maintenance plan.⁴⁹ However, assistance that is not reviewed and adjusted over time may become misaligned with current needs, and thereby less effective. Some papers also suggested that maintaining consistent and predictable assistance to recipient nations from year to year may diminish the influence the United States has with respect to that country. However, others suggested that initial influence itself is overestimated, and assistance continuity is irrelevant.

Development Assistance

There is only limited evidence on the effect of larger amounts of aid on growth. Dutta, Mukherjee, and Roy identified a positive and strong relationship between aid amount received and a country's investment.⁵⁰ A positive relationship between aid and growth was also identified by Clemens et al., who used instrumental variables and generalized method of moments.⁵¹ In addition to finding a positive relationship between quantity of aid and growth, however, these studies also identified diminishing returns to scale for the quantity of aid disbursed on economic growth and investment. Therefore, the positive impact of aid on growth decreases at higher levels of aid disbursed.

The amount of foreign aid disbursed in previous periods seems to be a determinant of future aid disbursements, but the evidence remains limited and results should be interpreted with caution. Fuchs, Dreher, and Nunnenkamp defined aid inertia as continuity in the amount of aid disbursed by donors to certain countries that had been receiving aid in previous periods of time.⁵² They used panel data with fixed effects to analyze the impact of aid on growth, controlling for aid inertia. They found that aid budgets evolve slowly over time, and thus donor generosity seems to continue in subsequent periods for those countries that have received aid support in the past. The effects identified for aid continuity as a determinant of future aid disbursements are particularly strong.

⁴⁹ Moroney, Thaler, and Hogler, 2013.

⁵⁰ Dutta, Mukherjee, and Roy, 2015.

⁵¹ Michael A. Clemens, Steven Radelet, Rikhil R. Bhavnani, and Samuel Bazzi, "Counting Chickens When They Hatch: The Short Term Effect of Aid on Growth," Center for Global Development, Working Paper 44, 2004; Michael A. Clemens, Steven Radelet, Rikhil R. Bhavnani, and Samuel Bazzi, "Counting Chickens When They Hatch: Timing and the Effects of Aid on Growth," *The Economic Journal*, Vol. 122, No. 561, 2012.

⁵² Fuchs, Dreher, and Nunnenkamp, 2014.

Conclusion

While the literature remains mixed on many aspects of the relationship between individual factors and the success of security cooperation activities, there are a few broad-stroke conclusions that seem well supported in both fields:

1. Countries that are considered strategically important receive more aid than countries that are not.
2. Countries with good governance records receive more aid, though governance is often conflated with stability rather than a measure of democratic principles.
3. Countries that previously received aid continue to receive aid.
4. Aid is used more effectively in countries with good governance.
5. Aid is more effective when it is given consistently over time, but there are diminishing returns on investment.

There was some evidence in both security cooperation and foreign aid to support the hypothesis that cultural ties between donors and aid recipients are a positive determinant of aid disbursement, and that more stable and lower-income countries tend to experience more immediate positive results after receiving aid. However, the evidence for these conclusions was more limited, and more evidence would be desirable before a strong conclusion can be made.

Undertaking a comparative literature review of security cooperation activities and outcomes and nonmilitary foreign assistance led to several additional findings. First, while nonmilitary foreign assistance has evolved in its theory and practice more substantially than security cooperation, it has evolved unevenly and the most rigorous studies remain in a smaller subset of functional areas.⁵³ Second, there are limited security cooperation studies that suggest a definite relationship between the kinds of countries that receive security cooperation aid and the kind of aid they receive. Third, there are even more limited studies on whether the aid that a country receives is used effectively, as defined by the security cooperation objectives agreed upon by the United States and the partner country. Fourth, the lack of definitive analysis in security cooperation is largely the result of continued data scarcity and requires analysts to think creatively about other ways to assess the effectiveness of security cooperation. Each point is important and worthy of further thought. Overall, however, this review highlights the paucity of data and analyses currently available to definitively guide Army planners in what types of activities have historically been most effective. The Army cannot rely on previous analyses as a substitute for effective AM&E going forward.

That nonmilitary foreign assistance does not provide the kind of definitive answer to the question of relationships between factors and outcomes that was initially sought should be thought of as less of a critique of development literature and as more of a hopeful forecast for analytic progress in both fields. As noted in earlier sections, assessing the effectiveness of nonmilitary foreign aid has gone through several periods of evolution, from viewing randomized control trials conducted in austere testing conditions as best practice to seeing other forms of evaluation as more preferable given the realities of project implementation and assessment.

⁵³ Most rigorous studies for who receives assistance were identified in the following areas: strategic importance, democratic institutions, and consistency of aid. Most rigorous studies for whether assistance is effective were identified in the following areas: quantity of aid and good governance.

Creating greater certainty in linking a given country characteristic, a given security cooperation activity or portfolio of activities, and the desired strategic objective or outcome requires the Army to think about new ways to collect and assess data. While having greater access to better quantitative data upon which more rigorous statistical models may be built is undoubtedly a step in the right direction, it is not the only solution. Gathering qualitative data on country conditions before and after project implementation, conducting project performance evaluations, and placing greater emphasis on requiring feedback in the form of AARs (or something similar) from stakeholders are all ways of getting at higher-quality data without relying exclusively on quantitative data.

In the next chapter, we examine how recent Army security cooperation activities match country characteristics associated with greater assistance effectiveness.

Army Security Cooperation Empirics

Previous studies in security cooperation and international development have explored how differences between countries correspond to differences in the volume of assistance rendered, and the extent to which it is effective. In this chapter, we examine how Army security cooperation in recent years matches recommendations from previous analyses. First, we examined activity attendance—which country characteristics correspond to increased chances of attending Army security cooperation activities. Second, we assessed the extent to which these characteristics correspond with what research considers most effective.

Our analysis confirms that countries with certain characteristics are especially likely to participate in security cooperation activities. In particular, the United States tends to favor stable democracies with well-funded militaries. However, we consistently found that our results were weaker than anticipated based on a survey of the literature, because different country characteristics corresponded to different kinds of assistance, rather than different volumes of assistance. For example, while country stability is a strong predictor of security cooperation in general, country *instability* is a very strong predictor of humanitarian assistance.

To understand the specificity of Army security cooperation, we examined the share of each country's engagement portfolio devoted to each of six broad goals of security cooperation:

1. Outreach—demonstrating goodwill and laying the groundwork for future relations. Examples: humanitarian assistance and transnational threat cooperation.
2. Professionalization—professionalization of local military and stabilization of the local security environment. Examples: human rights training, professional military education, and peacekeeping operations.
3. Operational effectiveness—improving the operational effectiveness of partner militaries. Examples: foreign military financing (FMF), foreign military sales (FMS), and advanced military education.
4. Shared values and priorities—facilitating convergence on mutually desired ends of security cooperation and acceptable ways of accomplishing those ends. Examples: political-military dialogue, senior leader engagement, efforts to build doctrinal consensus, and educational programs to increase support for norms such as democracy and civilian control of the military.
5. Joint operations—improving the ability of U.S. and partner militaries to work closely to accomplish common ends. Examples: interoperability, personnel exchanges, information-sharing, and research and development (R&D).

6. Deterrence and reassurance—discouraging potential adversaries from aggressive action, and improving the ability of the United States to win wars if deterrence fails. Examples: combat exercises, access agreements, surveillance of potential aggressors, and logistics.

These six broad goals are derived from previous RAND research into the activities through which the Army enacts security cooperation engagement. Appendix A provides more detail into how these categories were developed and how they relate to each other. We examined the security cooperation portfolios of over 150 countries to determine how much of each portfolio contributed to each of these six goals, and how differences in portfolio share corresponded to differences in country characteristics. We found that Army security cooperation is heavily tailored to match GCC objectives. Countries within a GCC tend to be more similar than the average two countries, so these objectives probably also reflect the country characteristics common to most countries in that GCC. However, within each GCC, differences in partner country characteristics consistently correspond to differences in portfolio share devoted to each of the goals.

Approach

To examine the relationship between Army security cooperation, partner country characteristics, and GCC objectives, we undertook statistical analyses of Army security cooperation. We examined over 9,000 Army security cooperation activities conducted between 2009 and 2014 and recorded in the DoD-wide Global-Theater Security Cooperation Management Information System (G-TSCMIS). G-TSCMIS is the DoD's data repository for security cooperation activities. All security cooperation activities are required to be included; however, based on previous RAND research, we believe there was not complete compliance with these reporting requirements.¹ As a result, although there are gaps in G-TSCMIS's data coverage, it is the most comprehensive dataset available for the examination of U.S. security cooperation activities. It can provide an illustrative overview of many of the Army's activities with partner countries.² It is important to note, however, that not all activities are created equally. While this analysis focuses on patterns in overall activity attendance, the impact of engaging with a partner in a two-day subject matter expert exchange is much smaller than holding a multinational joint exercise.

¹ See McNerney, 2016; and O'Mahony et al., 2017b. Although there are gaps in coverage, ASCC interviews did not highlight systematic biases in types of activities excluded from G-TSCMIS. There is an expectation that compliance with G-TSCMIS reporting improved over time. Based on our analysis of included activities, we believe compliance prior to 2009 was much lower than from 2009. For further sensitivity analyses of G-TSCMIS, see Joshua Mendelsohn, Elizabeth Bodine-Baron, Angela O'Mahony, and Thomas S. Szayna, *A Snapshot of Army Engagement with Partners: Global Landpower Network Baseline*, Santa Monica, Calif.: RAND Corporation, July 28, 2016. Not available to the general public.

² Much of the data included in G-TSCMIS are not commensurable, and many of the data fields are text based. To leverage these data we employed a big data analytic research approach and employed a latent space analysis to identify underlying dimensions in the data. Department of the Army, *Global-Theater Security Cooperation Management System (G-TSCMIS) Guidance and Business Rules*, PR 2456: GLN report, Washington, D.C.: Army Doctrine Reference Publication Headquarters, 2014b.

Data Used in Our Analyses

The Defense Security Cooperation Agency (DSCA) is responsible for maintaining and developing the G-TSCMIS database, and the specific entities sponsoring the activity are responsible for data entry. G-TSCMIS came online in 2010, but it builds on the GCC-level data collection infrastructure instituted in 1998–1999, and includes activity data from it. G-TSCMIS also pulls data from the Overseas Humanitarian Assistance Shared Information System.³

G-TSCMIS entries identify which countries participated in each security cooperation activity, as well as information about activity facilitators, activity title, descriptions, common activity types, and military contact themes (which describe the purposes of the activities). Our analyses draw heavily on information about which countries participated, what types of activities were undertaken, and what the purposes of the activities were. Although the military contact themes provide important information about the purpose of particular activities, they are also somewhat idiosyncratic. To examine broad categories of strategic objectives for Army security cooperation, we build on previous RAND categorizations of G-TSCMIS activities into six keyword-derived categories, which are discussed in detail in Appendix A.⁴

In the previous chapter, we divided country characteristics into five groups: strategic, political, absorptive capacity, cultural, and financial assistance.

Strategic considerations are country characteristics that may make a country especially critical to the United States' international goals. We examined three empirical measures of strategic considerations: the size of the partner country's economy (gross domestic product [GDP]) as a proxy for the country's potential impact on the world economy;⁵ countries' GDP per person as a proxy for their economic power;⁶ and support for U.S. positions during important UN votes as a proxy for countries' likelihood of using international influence to support the United States.⁷ After multiple rounds of preliminary analyses, we found that GDP per person and UN voting support most effectively measured the aspects of strategic importance that matter for activity attendance. Findings for those two are presented in this chapter. All of our model specifications are presented in Appendix B.

Political considerations are country characteristics that make the country's government a more desirable partner for the United States, either because the government is more effective at

³ G-TSCMIS contains activities up through the SECRET level of classification. However, these analyses focus on the 91 percent of events that are unclassified. We conducted sensitivity analysis to determine whether excluding classified information leads us to different conclusions. We found only a few cases where excluding the classified data led us to underestimate the true extent of the security cooperation relationship between the United States and the partner nation. These underestimations do not change the substance of our findings.

⁴ Since mission-level operators/users input data, there are variations in data quality and consistency among users, as well as GCC differences in the activity templates used. Moreover, due to differences in GCC data collection across GCC, we control for differences in GCC reporting in our analyses.

⁵ World Bank, *World DataBank*, 2016.

⁶ World Bank, 2016.

⁷ U.S. Department of State, *Voting Practices in the United Nations 2014: Report to Congress Submitted Pursuant to Public Laws 101-246 and 108-447*, July 2015. We measured countries' political alignment with the United States based on countries' shared foreign policy goals with the United States, as measured by the similarity in the country's voting record with that of the United States on votes at the UN that were designated as important by the U.S. Department of State. A country that always votes against the United States has a similarity score of 0. A country that always votes with the United States has a similarity score of 100. A country that votes with the United States one-half of the time and against the United States one-half of the times has a similarity score of 50.

achieving security or because partner governance is more consistent with U.S. values. We examined four empirical measures of political considerations: respect for human rights,⁸ democratic governance,⁹ state fragility,¹⁰ and number of recent coups.¹¹ After multiple rounds of preliminary analysis, we found that state fragility and democratic governance most effectively measured the aspects of political considerations that matter for activity attendance. Findings for those two are presented in this chapter. All of our model specifications are presented in Appendix B.

Absorptive capacity considerations are country and military characteristics that make the military better able to absorb new knowledge and resources from the United States, as well as characteristics that correspond to a more effective military. We examined four measures of absorptive capacity considerations: high school graduates¹² of military service age in the population, share of GDP devoted to military expenditures,¹³ military technical capability,¹⁴ and military professionalism.¹⁵ After multiple rounds of preliminary analysis, we found that high school graduates and GDP devoted to military expenditures most effectively measured the absorptive capacity features that matter for activity attendance. Findings for those two are presented in this chapter. All of our model specifications are presented in Appendix B.

Cultural considerations are aspects of a nation's history, way of life, and institutions that may be more or less similar to the United States, facilitating relationship-building. To measure cultural compatibility, we use the percentage of the population that speaks English,¹⁶ because having an English-speaking population corresponds to (a) the ability to communicate easily with the U.S. personnel, (b) a shared colonial history, replete with institutions that evolved from a common colonial template, and (c) the ability to be influenced by the same global media.

Financial considerations are aspects of the U.S. financial assistance relationship with the country. We examined two measures of financial assistance considerations: average dollar amounts of financial assistance¹⁷ and average annual variation in dollar amounts (normalized).¹⁸ These measure the magnitude and inconsistency of financial assistance, respectively.

⁸ David Cingranelli, David L. Richards, and K. Chad Clay, "The CIRI Human Rights Dataset," *CIRI Human Rights Data Project*, Version 2014.04.14, 2014.

⁹ Monty G. Marshall and Benjamin R. Cole, "Global Report 2014: Conflict, Governance, and State Fragility," *Center for Systemic Peace*, 2014.

¹⁰ Marshall and Cole, 2014.

¹¹ Marshall and Cole, 2014.

¹² World Bank, 2016.

¹³ We drew on previous RAND research to identify proxy variables for military capabilities and absorptive capacity. O'Mahony et al., 2017a.

¹⁴ IHS Jane's, *Sentinel Security Assessments*, 2015.

¹⁵ We drew on previous RAND research to identify proxy variables for military capabilities and absorptive capacity (see O'Mahony et al., 2015a).

¹⁶ We examined statistics on English speakers around the world. These data are available in multiple forms from multiple sources. After vetting the article's sources, we rely on the thorough compilation of statistics available on Wikipedia. Wikipedia, "List of Countries by English-Speaking Population," n.d.

¹⁷ U.S. Agency for International Development, *U.S. Overseas Loans and Grants: Obligations and Loan Authorizations, July 1, 1945–September 30, 2014*, n.d.

¹⁸ U.S. Agency for International Development, n.d.

Approach to Statistical Analysis

We use logistic regression to calculate how past security cooperation engagement differs among countries, based on the characteristics of those countries. Logistic regression modeling assumes that each country characteristic makes the chances of security cooperation engagement more or less likely, and then attempts to calculate how much more or less likely, based on the differences between countries that have more of that characteristic and countries that have less. For our first round of modeling, we look at over 9,000 engagement activities conducted between 2009 and 2014 and recorded in G-TSCMIS, and calculate the chances that a particular country attended each of them, given the characteristics of that country. To interpret our results about partners' attendance at security cooperation activities, we ask the question, "For every 1,000 activities reported to G-TSCMIS, how many did country X attend?" For our second round of modeling, we categorize G-TSCMIS activities into six overlapping types, and calculate the chances that each of a country's attended activities falls into that category. This is mathematically equivalent to calculating the proportion of that country's engagement portfolio that falls into that engagement category. To interpret our results about the United States' security cooperation portfolio with each country, we ask the question, "Out of all the activities that country X attended, what percentage could be categorized as having type Y aspects?" In this chapter we discuss the results and implications of these analyses. Appendix B reports the actual modeling coefficient tables calculated, as well as more technical details of our statistical approach.

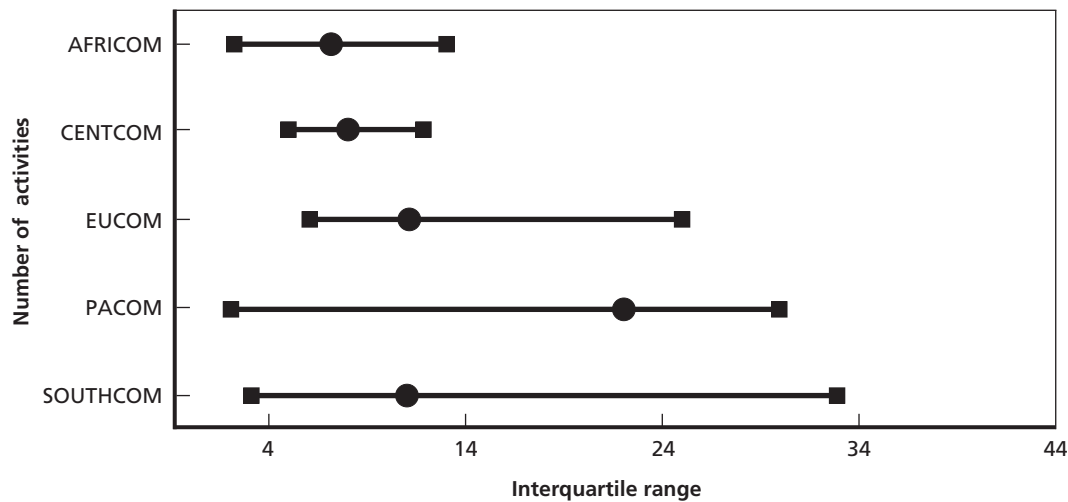
Overview

Countries differ in the intensity with which the Army engages them in security cooperation activities. The average U.S. security partner attends 12 out of every 1,000 security cooperation activities. This average varies across GCCs. On the low end, the average United States Africa Command (AFRICOM) partner attends 7 per 1,000, and the average United States Central Command (CENTCOM) partner attends 8. At the high end, the average United States Pacific Command (PACOM) partner attends 22 activities per 1,000. Within each GCC, countries vary widely in their average rates of activity attendance as well.

Figure 3.1 depicts the range of activity attendance rates within each GCC. The x-axis measures the number of activities a country attends, per 1,000 security cooperative activities conducted across all U.S. security partners. The dot indicates the average rate of attendance for each GCC and responds to the figures listed above. The small squares indicate the twenty-fifth and seventy-fifth percentiles of activity attendance.

Different GCCs show different approaches to engagement. In the United States European Command (EUCOM) and the United States Southern Command (SOUTHCOM), the Army has strong relationships with most countries, and extraordinarily close relationships with a few. Thus, the GCC average attendance rate dots are close to the global average, but seventy-fifth-percentile squares are more than double the average. In AFRICOM and CENTCOM, the Army engages in security cooperation with most countries, but that engagement is much weaker than that of SOUTHCOM or EUCOM. The Army's stronger partners (seventy-fifth percentile) in AFRICOM and CENTCOM have activity participation rates only slightly higher than the average EUCOM/SOUTHCOM partner. PACOM

Figure 3.1
Interquartile Range of Activity Attendance Rates by GCC



SOURCE: RAND analysis.

NOTE: This figure plots the interquartile range of activity attendance rates—the gap between the twenty-fifth percentile and the seventy-fifth percentile for the number of activities attended per 1,000. The circle in the middle of each range denotes the average country.

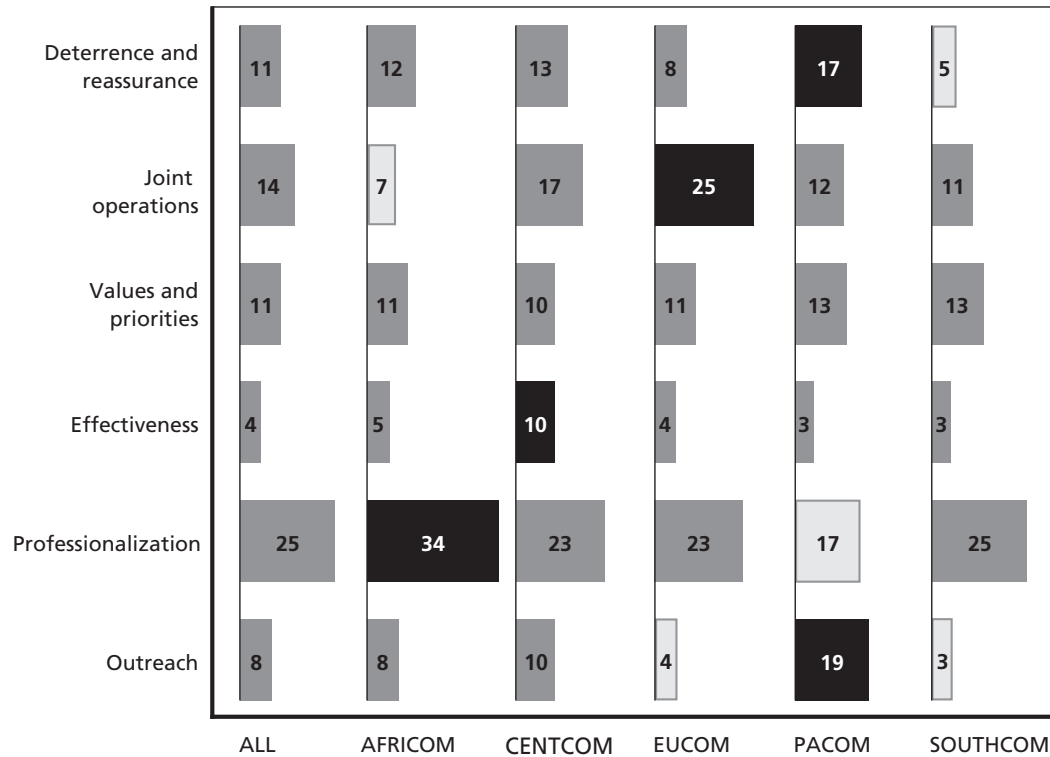
RAND RR2165A-3.1

participation rates are the highest by far. Renewed policy emphasis on the region during our data collection window, the presence of some of the United States' most long-standing security challenges, the presence of some of the United States' strongest alliances, and the proliferation of high-capability militaries in the region all provided strong incentives for the Army to engage deeply with a large roster of countries. Consequently, PACOM is the only GCC¹⁹ where the average engagement rate is closer to the seventy-fifth percentile than to the twenty-fifth.

GCCs also differ in their balance of security cooperation objectives. Figure 3.2 reports the average share of a country's security cooperation activity portfolio devoted to each of the six objectives (outreach, professionalization, operational effectiveness, shared values and priorities, joint operations, and deterrence and reassurance) for each GCC. Globally, *professionalization* generally accounts for more portfolio share than any other category (25 percent for a typical country), while *operational effectiveness* generally accounts for the least (4 percent for a typical country). In contrast, AFRICOM countries tend to have more portfolio share involving *professionalization*, and less involving *joint operations*, because of the relatively low quality of the typical military in the region. CENTCOM countries tend to have more *operational effectiveness* because of the high volume of weapon sales and larger than typical national budgets with which to purchase them. EUCOM countries tend to have more *joint operations* and less *outreach* because of the presence of many NATO allies. PACOM tends to have more *deterrence and reassurance* and *outreach* because of elevated risks posted by countries in the area of responsibility with a tenuous U.S. relationship, while SOUTHCOM tends to have less *deterrence and reassurance* and *outreach* because of the relatively low risks.

¹⁹ Of the five displayed. U.S. Northern Command has a similar profile.

Figure 3.2
Average Portfolio Share Percentage for Each Activity Type in Each GCC



SOURCE: RAND analysis of G-TSCMIS.

KEY: Black indicates the portfolio share percentage for that GCC is more than three percentage points higher than the average (The ALL column). White indicates the portfolio share percentage for that GCC is more than three percentage points lower than the average.

RAND RR2165A-3.2

Strategic Considerations

One reason for differences in activity attendance rates is that countries vary in their relative importance for accomplishing U.S. interests. Our review of previous analyses found that political and economic strategic considerations have been associated with high levels of assistance in both the security and development realms. We examine both a measure of political alignment with the United States (captured by countries' supportiveness of U.S. positions on important UN votes) and a measure of economic capabilities (captured by GDP per capita) to determine whether Army security cooperation favors countries that are strategically important to the United States—that is, countries that are more closely aligned with the United States, or have greater GDP per capita.²⁰ In contrast with previous analyses, our statistical analysis found that, on average, recent Army security cooperation activities have been held disproportionately with countries with lower GDP per capita and lower political

²⁰ Lauren Ploch, "Africa Command: U.S. Strategic Interests and the Role of the U.S. Military in Africa," *Congressional Research Services*, July 2011.

alignment with the United States (as captured by concurrence in partners' and U.S. voting behavior at the UN).

While this result appears surprising at first blush, we expect this reflects the different types and purposes of security cooperation for different countries. As we discuss in more detail below, countries that are more politically aligned with the United States are more likely to participate in security cooperation activities that build capacity for joint operations, and countries with higher GDP per capita are more likely to be engaged through activities that build military effectiveness. These types of activities tend to be the largest, most-resource-intensive activities the Army undertakes. In contrast, countries that have lower GDP per capita and are less politically aligned with the United States are more likely to be engaged through activities that provide humanitarian assistance, demonstrate the value of closer U.S. relations in a nonthreatening manner, and encourage convergence toward shared values and priorities—activities that can potentially reduce partners' vulnerabilities and strengthen relations with the United States.

Taken together, we find that Army security cooperation practices reflect a nuanced response to differences in strategic importance. While our literature review led us to expect that strategically important countries receive more aid, we found that different kinds of strategic importance correspond with different strategies of engagement.

Below, we present detailed analyses of the relationship between countries' strategic importance and their attendance in Army security cooperation activities.

Activity Attendance Based on Countries' Strategic Importance

Our statistical analysis (reported in Table B.1 in Appendix B) found negative and statistically significant correlations between activity attendance and a partner country's level of GDP per capita and its political alignment with the United States. Countries with *lower* GDP per capita and *lower* political alignment with the United States tended to participate in security cooperation activities at higher rates than countries with *higher* GDP per capita and political alignment with the United States.

What does this mean in terms of countries' event participation? Our statistical analysis found that countries that score in the ninety-fifth percentile of their GCC for GDP per capita, on average, were predicted to attend 1–3 fewer activities per 1,000 activities held than countries scoring in their respective GCC's fifth percentile.²¹ Countries that scored in the ninety-fifth percentile for their GCC for supporting U.S. positions in the UN (high political alignment with the United States) were predicted to attend 2–11 fewer activities per 1,000 activities on average than countries that scored in the fifth percentile. Taken together, an otherwise average country that, compared with other countries in its GCC, has a low GDP per capita (fifth percentile) and is not aligned politically with the United States (fifth percentile) could expect to attend 12 activities per 1,000. In contrast, a country with high GDP per capita (ninety-fifth percentile) and greater political alignment with the United States (ninety-fifth percentile) with the same characteristics could expect to attend 8 out of every 1,000 activities.

²¹ We predict the chances that a country will participate in any given security cooperation activity, and then multiply that chance percentage by 1,000 to get the number of activities that might involve the country, per 1,000 security cooperation activities. For example, if a country has a 1 percent chance of attending any given security cooperation activity, then $1\% \times 1,000 = 10$ activities.

This represents a 50 percent increase in activity participation. These comparisons are presented in Table 3.1.²²

This predicted outcome is driven more strongly by low political alignment than low GDP per capita. As can be seen in Table 3.1, holding the level of political alignment constant, increasing GDP per capita from low to high results in a predicted increase of one activity. In contrast, holding the level of GDP per capita constant, increasing political alignment from low to high results in a predicted increase of three activities.

Portfolio Share Based on Countries' Strategic Importance

Comparing the security cooperation portfolios for countries across our two strategic importance dimensions (political alignment with the United States and GDP per capita), differences in the weight given to different strategic objectives stand out.²³ Table 3.2 reports how the percentage of a country's security cooperation portfolio devoted to each of six security cooperative goals tended to differ, depending on the strategic importance of that country.

Table 3.1
Predicted Activity Attendance, as Strategic Importance Varies

	High GDP per Capita	Low GDP per Capita
High political alignment with United States	8	9
Low political alignment with United States	11	12

SOURCE: RAND analysis.

NOTE: Each table cell reports the expected number of security cooperation activities that a country might attend, out of every 1,000 security cooperation activities held, holding all other characteristics equal.

Table 3.2
Predicted Difference in Portfolio Share, Given Strategic Importance

	Outreach	Professionalization	Operational Effectiveness	Values and Priorities	Joint Operations	Deterrence and Reassurance
GDP per capita	↑	↓	↑	↑	↓	↑
Political alignment with United States	↓				↑	↓

SOURCE: RAND analysis of G-TSCMIS.

NOTE: An up arrow indicates that having more of a strategic characteristic corresponded to having more portfolio share devoted to a specific security cooperative goal. A down arrow indicates that having the characteristic corresponded to having less portfolio devoted to the security cooperative goal. A black arrow signifies that the result is statistically different from a null effect at the 95 percent level. A gray arrow signifies that the result is statistically different from a null effect at the 90 percent level.

²² Table 3.1 presents a comparison of how activity attendance is expected to vary across different levels of GDP per capita and political alignment. The table reports the number of security cooperation activities that an otherwise average country could expect to attend, given its strategic characteristics. The activities are reported in increments of activities per 1,000 total security cooperation activities. For example, a value of 10 would mean that, for every 1,000 security cooperation activities held anywhere around the globe, a country with those characteristics might expect to attend 10 of them.

²³ The models underlying this sections are in Table B.2 in Appendix B.

Based on these results, countries' GDP per capita appears much more highly correlated with specific security cooperation objectives than political alignment with the United States.²⁴ Countries with higher GDP per capita are more likely to participate in activities focused on outreach, operational effectiveness, values and priorities, and deterrence and reassurance, and are less likely to participate in activities focused on professionalization and joint operations. Countries that are strongly aligned with the United States are more likely to take part in joint operations, and less likely to take part in outreach and deterrence and reassurance activities. Of particular note, countries that are not closely aligned with the United States but have higher GDP per capita are disproportionately likely to participate in outreach and deterrence and reassurance activities. Taken together, these activities are often used to develop nascent relationships in strategically important regions. In contrast, countries that are closely aligned with the United States but have lower GDP per capita are disproportionately likely to participate in joint operations activities. These are often used to plan and deconflict operations with regional partners.

Political Considerations

Our review of previous studies found that countries with stronger and more democratic institutions were more likely to both receive international assistance and use assistance more effectively. We examine countries' democratic governance and fragility to determine whether Army security cooperation favors countries that are more democratic and politically stable.

Across our analyses of the relationship between political considerations and countries' participation in security cooperation activities, we found that, on average, Army security cooperation practices tend to favor stable democracies, which is consistent with the general thrust of studies in our literature review. That said, we also found that the relationship between partners' political characteristics and participation in security cooperation activities is nuanced. For example, the literature finds that during the Cold War, a time of heightened military threat, assistance was less likely to favor stable democracies. Consistent with this, we found that deterrence-oriented assistance does not have a strong statistical association with political characteristics. Likewise, the literature points out that assistance is more likely and more impactful among countries that are transitioning away from authoritarian governance or emerging from conflict (i.e., unstable). Consistent with this, we found that outreach-oriented assistance generally has a higher portfolio share among less stable, authoritarian countries, and operational effectiveness-oriented assistance has a higher portfolio share among less stable, democratic countries. We believe that one of the key drivers of the overall statistical relationship is that professionalization and joint operations-oriented assistance have larger portfolio shares among stable democracies; these types of activities are the largest ones by percentage share of the global security cooperation portfolio.

Below, we present detailed analyses of the relationship between countries' political considerations and their attendance in Army security cooperation activities.

²⁴ We believe the low correlation between security cooperation and type of security cooperation objective may reflect the relatively low political alignment scores for U.S. partners in CENTCOM and reflect long-standing differences at the UN between the United States and Middle East countries regarding Israeli and Palestinian issues.

Activity Attendance Based on Country Political Considerations

Our statistical analysis found positive and statistically significant correlations between activity attendance and a partner country's level of democracy and domestic stability.²⁵ More democratic and more stable partners tended to participate in security cooperation activities at higher rates than less democratic and less stable partners. Translating these results into numbers of activities, our statistical analysis found that countries that score in the ninety-fifth percentile of their GCC for state stability (very fragile) attended, on average, 4–13 more activities per 1,000 activities, compared with countries scoring in their GCC's fifth percentile. Countries that score in the ninety-fifth percentile for democratic governance (very democratic) tended to participate in 2–8 more activities per 1,000 activities, compared with countries scoring in their GCC's fifth percentile.

Taken together, an otherwise average country that, compared with other countries in its GCC, is highly stable and democratic could expect to attend 13 activities per 1,000. In contrast, a nondemocratic, unstable country could expect to attend 7 out of every 1,000 activities. Thus, a country with a stable, democratic governance is predicted to attend 38 percent more activities, compared with a country without those governance characteristics. As can be seen in Table 3.3, these differences in predicted activity attendance are driven more strongly by variations in domestic stability than in quality of democracy.

Portfolio Share Based on Country Political Considerations

Table 3.4 reports how the percentage of a country's security cooperation portfolio devoted to each of six security cooperative goals tended to differ, depending on the partner country's domestic political environment.²⁶

Three main observations emerge from Table 3.4. First, there is a strong positive correlation between countries' stability and quality of democracy and their participation in joint operations activities. This relationship captures the large focus in Army security cooperation on interoperability with key partners.

Second, there is a strong negative correlation between countries' stability and quality of democracy and their participation in outreach activities—outreach activities are disproportionately likely to be held with less democratic or less stable countries. Based on the activities included in our analyses, for unstable countries, outreach activities often consist of emergency assistance to prevent an insecure situation from deteriorating further. For nondemocratic

Table 3.3
Predicted Activity Attendance, as Political Characteristics Vary

	Unstable	Stable
Democratic	8	13
Authoritarian	7	11

SOURCE: RAND analysis.

NOTE: Each table cell reports the expected number of security cooperation activities that a country might attend, out of every 1,000 security cooperation activities held, holding all other characteristics equal.

²⁵ The model for this analysis is reported in Table B.1 in Appendix B.

²⁶ The models underlying this section are in Table B.2 in Appendix B.

Table 3.4
Predicted Difference in Portfolio Share, as Political Characteristics Vary

	Outreach	Professionalization	Operational Effectiveness	Values and Priorities	Joint Operations	Deterrence and Reassurance
Stable	↓	↑	↓		↑	
Democratic	↓		↑		↑	↑

SOURCE: RAND analysis.

NOTE: An up arrow indicates that having more of a strategic characteristic corresponded to having more portfolio share devoted to a specific security cooperative goal. A down arrow indicates that having the characteristic corresponded to having less portfolio devoted to the security cooperative goal. A black arrow signifies that the result is statistically different from a null effect at the 95 percent level. A gray arrow signifies that the result is statistically different from a null effect at the 90 percent level.

countries, especially nondemocratic countries with significant military capability, outreach activities represent a nonthreatening way of demonstrating the benefits of working *with* the United States, instead of working counter to American interests.

Third, operational effectiveness–oriented activities tended to be a higher share of a country’s engagement portfolio when that country is democratic but faces more domestic instability. We believe this pattern may be partly endogenous, because countries with security challenges are more likely to need military equipment and financing. It does, however, suggest that the United States is more inclined to offer such assistance to democratic regimes.

Absorptive Capacity Considerations

Our review of previous studies found that countries with greater absorptive capacity were more likely to receive international assistance and to use assistance more effectively.²⁷ However, this can be difficult to test because so many different factors can potentially fall under the label of absorptive capacity.²⁸ We examine a measure of knowledge transfer capability (population percentage of high school graduates²⁹) and a measure of military resources (military spending as a percentage of GDP) to determine whether Army security cooperation favors militaries with stronger absorptive capacity. Using military spending and population education levels as proxies for absorptive capacity, we find that the Army is far more likely to conduct security cooperation activities with capable partners.

Below, we present detailed analyses of the relationship between countries’ absorptive capacity and their attendance in Army security cooperation activities.

Activity Attendance Based on Absorptive Capacity Considerations

Our statistical analysis found positive and statistically significant correlations between activity attendance and a partner country’s level of education and military spending.³⁰ Countries with

²⁷ Paul et al., 2013b.

²⁸ Lamb and Mixon, 2013.

²⁹ While military aid research on the effects of education are limited, research on foreign direct investment (FDI) supports the notion that education is a significant enabler of absorptive capacity. Miao Wang and M. C. Sunny Wong, “FDI, Education, and Economic Growth: Quality Matters!” *Atlantic Economic Journal*, Vol. 39, No. 2, 2011.

³⁰ The model for this analysis is reported in Table B.1 in Appendix B.

militaries that are amply funded and that can draw from educated populations attend security cooperation activities at a higher average rate. Countries that score in the ninety-fifth percentile of their GCC for proportion of high school graduates in the population tended to participate in 2–8 more activities per 1,000 activities, compared with countries scoring in their respective GCC's fifth percentile. Countries that score in the ninety-fifth percentile for military spending as a proportion of GDP (promotes absorption) tended to participate in 7–29 more activities per 1,000 activities. An otherwise average country that, compared with other countries in its GCC, has an educated population (ninety-fifth percentile) and a large share of GDP devoted to military spending (ninety-fifth percentile) could expect to attend 16 of every 1,000 activities. In contrast, its antithesis could expect to attend 5 of every 1,000 activities.

Table 3.5 displays our linear modeling projections for how the predicted rate of activity attendance might change, depending on the absorptive capacity characteristics of a country. Both greater military spending and a population rich in high school graduates predict higher rates of activity attendance, but military spending is a much more powerful predictor than population education (7–9 additional activities vs. 2–3). Taken together, a country with many high school graduates and greater military spending is predicted to attend 220 percent more activities than a country without those governance characteristics.

Portfolio Share Based on Absorptive Capacity Considerations

Table 3.6 reports how the percentage of a country's security cooperation portfolio devoted to each of six security cooperative goals tended to differ, depending on the partner country's absorptive capacity.³¹

Table 3.5
Predicted Activity Attendance, as Absorptive Capacity Characteristics Vary

	Greater Military Spending	Less Military Spending
Many high school graduates	16	7
Few high school graduates	13	5

SOURCE: RAND analysis.

NOTE: Each table cell reports the expected number of security cooperation activities that a country might attend, out of every 1,000 security cooperation activities held, holding all other characteristics equal.

Table 3.6
Predicted Difference in Portfolio Share, as Absorptive Capacity Characteristics Vary

	Outreach	Professionalization	Operational Effectiveness	Values and Priorities	Joint Operations	Deterrence and Reassurance
High school graduates	↓	↑		↓		↑
Military spending	↓	↓	↑		↑	↑

SOURCE: RAND analysis of G-TSCMIS.

NOTE: An up arrow indicates that having more of a strategic characteristic corresponded to having more portfolio share devoted to a specific security cooperative goal. A down arrow indicates that having the characteristic corresponded to having less portfolio devoted to the security cooperative goal. Reported results are statistically different from a null effect at the 95 percent level.

³¹ The models underlying this section are in Table B.2 in Appendix B.

Based on the results in Table 3.6, partners' absorptive capacity—in both the long and shortterm—do appear to shape the types of activities that partners participate in. While both educational attainment and military expenditure may influence partners' absorptive capacity, military expenditure appears to have a larger impact on short-term, operational-level objectives. Partners with greater military expenditures are more likely to participate in activities focused on operational effectiveness and joint operations. In contrast, educational attainment is correlated with participation in the longer-time-horizon activities associated with professionalization.

Drawing on absorptive capacity more broadly, deterrence and reassurance activities tend to occupy more portfolio share for countries that provide their militaries with abundant funding and/or that have a well-educated population. These activities match well with partners for which the military will be well funded enough to maintain a strong deterrent, and that have well-educated personnel who can better absorb knowledge transfer from the United States.

Conversely, outreach activities tend to occupy more portfolio share for countries that have both low levels of population education and low levels of military resources. This may reflect the greater share of activities such as humanitarian assistance with partners that have low educational attainment and/or poorly resourced militaries.

Cultural Considerations

Army security cooperation engagement may also vary among countries based on their cultural similarity with the United States, including common language, customs, institutions, media, and colonial history.³² In this study, we examine language similarity (percentage of English speakers) for two reasons. First, the ability to communicate is among the most important aspects of cultural compatibility, as it can severely limit the ability of U.S. and partner military personnel to interact. Second, common language correlates strongly with other aspects of cultural similarity, such as shared historical experiences, shared media consumption, expatriate communities, and interpersonal ties.

Both our analysis and previous literature found that culturally similar countries are more likely to be involved in security cooperation activities. That said, while our results found a statistically significant effect for greater participation with English-speaking partners, the overall effect was substantively small. Our analysis of security cooperation portfolios provides more nuance on this effect, finding that the correlation between shared language and activity participation pertains most strongly to joint operations-oriented activities. Joint operations require both the closest levels of coordination and the highest levels of trust, so common language could be either a marker for the ability of military units to coordinate closely without language barriers or an indicator of long-standing relationships that date back to a shared colonial history.

Below, we present detailed analyses of the relationship between countries' cultural considerations and their attendance in Army security cooperation activities.

Activity Attendance Based on Cultural Considerations

Our statistical analysis found a positive and statistically significant correlation between activity attendance and a partner country's proportion of English speakers in the population.³³ Coun-

³² Abbe, Gulick, and Herman, 2007.

³³ The model for this analysis is reported in Table B.1 in Appendix B.

Table 3.7
Predicted Difference in Portfolio Share, as Cultural Characteristics Vary

	Outreach	Professionalization	Operational Effectiveness	Values and Priorities	Joint Operations	Deterrence and Reassurance
English speakers	↓			↓	↑	↓

SOURCE: RAND analysis of G-TSCMIS.

NOTE: An up arrow indicates that having more of a strategic characteristic corresponded to having more portfolio share devoted to a specific security cooperative goal. A down arrow indicates that having the characteristic corresponded to having less portfolio devoted to the security cooperative goal. Reported results are statistically different from a null effect at the 95 percent level.

tries that score in the ninety-fifth percentile for proportion of English speakers in the population attended up to 6 more activities per 1,000 activities. An otherwise average country that, compared with other countries in its GCC, has an English-speaking population (ninety-fifth percentile) could expect to attend 10 of every 1,000 activities. In contrast, its antithesis could expect to attend 9 of every 1,000 activities. In practice, this effect is substantively quite small.

Portfolio Share Based on Cultural Considerations

Table 3.7 reports how the percentage of a country's security cooperation portfolio devoted to each of six security cooperative goals tended to differ, depending on the partner country's proportion of English speakers in the population.³⁴

Joint operations activities tend to be a larger share of a country's engagement portfolio when a large portion of that country's population speaks English. Previous RAND research has noted a shortage of language and cultural capacity within the Army. This can make it difficult to work closely with non-English speakers, and may cause joint operations engagement to favor countries with cultural and linguistic similarity.³⁵ In contrast, outreach and shared values and priorities activities tend to be a larger share of a country's engagement portfolio when a smaller portion of that country's population speaks English.³⁶

Financial Aid Considerations

Army security cooperation engagement may also vary among countries because of different financial assistance patterns between the United States and partner nations. We examine the average amount of financial assistance and the average annual variation in financial assistance to determine whether partner countries that receive large³⁷ and/or inconsistent³⁸ aid packages attend more security cooperation activities, as previous research would suggest.

³⁴ The models underlying this section are in Table B.2 in Appendix B.

³⁵ O'Mahony et al., 2017b.

³⁶ There is also a statistically significant negative correlation between the share of deterrence and assurance activities in a partner country's portfolio of security cooperation activities and its proportion of English speakers in the population. However, in further analysis at the GCC level, this finding did not appear to hold in all GCCs.

³⁷ Sislin, 1994.

³⁸ Clemens et al., 2012.

In agreement with previous scholarship, we found that financial aid is highly consistent from year to year—countries tended to receive the same amount and type of aid over time. As a result, neither the quantity of aid nor the relatively low level of variation in aid over time is strongly correlated with variation in the amount or type of aid a partner receives. We present our models for aid quantity and consistency in Appendix B, but do not present them in detail in this chapter.

Conclusion

There are definite patterns in which countries are more engaged through security cooperation activities. Overall, we found that Army security cooperation generally favors three types of countries: those in need of greater engagement (e.g., countries with high domestic instability), those with which the United States would like to improve relations (e.g., countries with low political alignment with the United States), and those with which greater engagement will be most productive (e.g., stable democracies with strong military capability).

However, it is perhaps more informative to say that different kinds of engagement are administered to different kinds of countries, based on perceptions of the problem to be solved and theories of change (TOCs) about what can be accomplished through specific types of engagement. For example, weapon sales and increased financing tend to be standard responses for democracies facing domestic instability. This implies a problem—a democratic government is at risk, and it is in the best interest of the United States to preserve it. It also implies a TOC—if the government could better arm and finance its military, it could better stabilize its domestic security situation.

Examining the engagement strategies currently in use to determine how effective they are at achieving desired end states requires following an established process of AM&E. At its heart, AM&E requires security cooperation planners to examine the assumptions they hold about how security cooperation activities connect to desired results. While development scholars and practitioners have been engaged in creating TOCs for international development assistance activities for the better part of two decades, these processes are relative newcomers to the Army's way of thinking about and planning for security cooperation activities. Part II of this report discusses what good AM&E looks like, what it does, and how it has been and can be further applied to Army activities.

PART II

Implementing an Army AM&E Framework

AM&E Developments and Implications for the Army

As Part I of this study discussed, Army security cooperation activities have aligned fairly well with what previous analyses have found contribute to effectiveness in both security cooperation and international development assistance. However, across the U.S. government, requirements for more rigorous and systematic AM&E are growing. The Army, with the rest of DoD, must improve its ability to analyze and report the extent to which its security cooperation activities are meeting their objectives. As Part I makes clear, there is still much to be learned. As the United States increases the scope of security cooperation, in terms of both activity and partner types, planners will need to understand not only what has been effective in the past but also how to identify lessons from the broader range of partners the United States is engaging. Part II of this study focuses on helping the Army develop an AM&E implementation framework that will better link security cooperation activities to their outcomes.

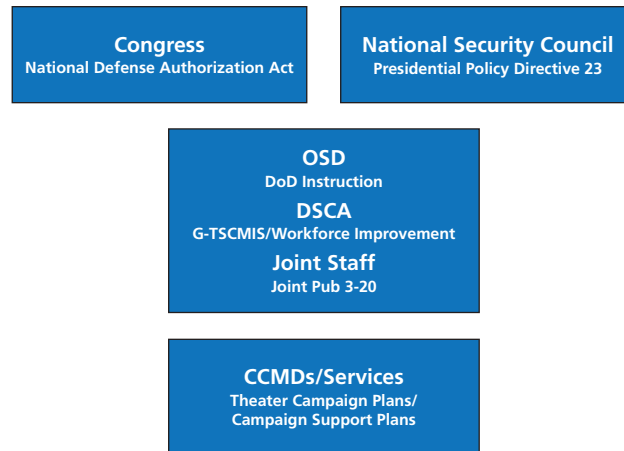
AM&E Developments—and Requirements—Are Accelerating

Within the past ten years, several developments have laid the foundations for a more strategic and analytically rigorous approach to security cooperation. As shown in Figure 4.1, these developments have arisen at several levels of government: Congress, the White House, OSD, DSCA, and combatant commands (CCMDs), and within the Army itself.

First, an important impetus for DoD's recent efforts to improve the security cooperation AM&E process has come from Congress, some of whose members are eager for the department to improve its ability to track security cooperation resource expenditures and demonstrate the effectiveness of security cooperation programs in achieving DoD objectives at the national, theater, or country level. In Section 1202 of the Fiscal Year (FY) 2016 National Defense Authorization Act (NDAA), Congress mandated that DoD work with the Department of State (DoS) to establish a strategic framework for security cooperation "to guide prioritization of resources and activities." Furthermore, it specified that one of the elements of this strategic framework would be "a methodology for assessing the effectiveness of Department of Defense security cooperation programs in making progress toward achieving the primary objectives, priorities, and desired end-states . . . including an identification of key benchmarks for such progress."¹ The FY 2017 NDAA reiterates the previous NDAA's call for a security

¹ U.S. Congress, *National Defense Authorization Act for Fiscal Year 2016*, n.d.a.

Figure 4.1
New Developments and Requirements Affecting AM&E



RAND RR2165A-4.1

cooperation AM&E framework, but goes further by requiring an “independent assessment of defense security cooperation programs.”²

Second, the National Security Council laid the groundwork in 2013 for a more integrated approach to security sector assistance. Directed primarily at DoS and DoD, Presidential Policy Directive 23 (Security Sector Assistance Policy) called for “a deliberate and inclusive whole-of-government process that ensures alignment of activities and resources with [U.S.] national security priorities.” Among other things, this meant ensuring greater consistency between programs and objectives, fostering transparency and coordination across U.S. government agencies, building sustainable partner security capacity through comprehensive sector strategies, being more selective in the use of security assistance resources to achieve the greatest impact, and informing policy with rigorous analysis, assessment, and evaluations. According to the directive, this latter effort would involve introducing “common standards and expectations for assessing security cooperation requirements” as well as “investing in monitoring and evaluation of security sector assistance programs.” Although the emphasis was on greater standardization of AM&E policies and procedures throughout the security assistance sector, the directive did indicate that newly established “standards and data collection will take into account the varying security and information environments where U.S. programs operate.”³

Third, in 2016 OSD released Department of Defense Instruction (DoDI) *Assessment, Monitoring and Evaluation Policy for the Security Cooperation Enterprise*, which provides guidance for a “whole of DoD” approach to security cooperation AM&E. It asserts that “accountability and learning are the primary purposes” of AM&E, and highlights AM&E’s importance in demonstrating security cooperation’s “returns on investment” so policymakers might “identify and improve or eliminate ineffective initiatives” as well as garner evidentiary support for policy and legislative proposals.⁴

² U.S. Congress, *National Defense Authorization Act for Fiscal Year 2017*, n.d.b, sections 1204 and 1205.

³ The White House, *White House Fact Sheet on U.S. Security Sector Assistance Policy*, 2013.

⁴ Office of the Secretary of Defense, 2016.

The DoDI specifies two main military service requirements in supporting CCMD AM&E: (1) an initial (country-level) assessment of the security cooperation environment and (2) an initial design document (IDD) for “significant” security cooperation initiatives. “The initial assessment provides an understanding of the context, conditions, partner capabilities, and requirements immediately before the implementation of security cooperation initiatives and other activities.” It informs initiative design and establishes a “baseline against which to track progress.” The IDD contains milestones and measures that allow for performance monitoring and “independent and rigorous evaluations” of initiative “relevance, effectiveness, and sustainability.”⁵ The IDD must also include a TOC.

Within the terms of the DoDI, the GCCs have the primary security cooperation AM&E role. Their responsibilities include (1) leading initial assessment efforts, (2) orchestrating the assessment process, (3) formulating IDDs for all significant initiatives, (4) monitoring all significant initiatives, and (5) submitting to DSCA all initial country assessments and IDDs for new security cooperation initiatives.⁶ Based on the analysis in this report, services play an important role in every step of the AM&E process, particularly for the activities they implement directly. ASCCs play a particularly important role linking Army security cooperation to GCCs’ initial assessment and IDD.

In addition to issuing the DoDI, OSD has encouraged the development of AM&E processes, for example, by highlighting the importance of AM&E in various other guidance documents and hosting DoD-wide and interagency workshops to share best practices. These efforts were accelerated and better integrated after OSD’s establishment of a Deputy Assistant Secretary of Defense for Security Cooperation in 2015. This office undertook several initiatives over the past two years, supported by RAND’s National Defense Research Institute, which focused on institutionalizing best practices in security cooperation planning, transparency, and AM&E.⁷

Fourth, innovative concepts for using technology to improve data collection have allowed hundreds of U.S. security cooperation practitioners around the globe to report on their activities in more methodical and comprehensive ways and for policymakers and program managers to provide more rigorous oversight. The creation of security cooperation management information systems by the GCCs and other DoD organizations provided unprecedented platforms for planning, resourcing, assessing, monitoring, and evaluating future security cooperation AM&E efforts. These largely independent initiatives have begun to coalesce in the form of a comprehensive tool for tracking DoD security cooperation activities and resources: G-TSCMIS, whose development is being managed by DSCA (with support from the Joint Staff) and is required to be used by all DoD security cooperation activity managers. G-TSCMIS is still a work in progress, and CCMDs face challenges getting all relevant activities included in the system.⁸ Still, even when G-TSCMIS is fully developed and widely used, it will not fulfill all the information management needs required for effective security cooperation planning and

⁵ Office of the Secretary of Defense, 2016.

⁶ Office of the Secretary of Defense, 2016.

⁷ See McNerney et al., 2016; and Jefferson P. Marquis, Michael J. McNerney, S. Rebecca Zimmerman, Merrie Archer, Jeremy Boback, and David Stebbins, *Developing an Assessment, Monitoring, and Evaluation Framework for U.S. Department of Defense Security Cooperation*, Santa Monica, Calif.: RAND Corporation, RR-1611-OSD, 2016.

⁸ O’Mahony et al., 2017b; Watts et al., 2017.

AM&E, which is why the CCMDs and other security cooperation organizations continue to use their own systems.

In addition to overseeing the development of G-TSCMIS, DSCA has also recently launched an initiative to formalize security cooperation workforce competencies and associated experiential and training requirements, which could eventually contribute to a larger pool of qualified AM&E professionals within the security cooperation field.

Fifth, the Joint Staff has also played an important role in attempting to improve and standardize security cooperation and security force assistance (SFA) doctrine, including with respect to AM&E. Acknowledging the challenge of determining the extent to which security cooperation activities have contributed to U.S. objectives, the draft *Joint Publication 3-20* provides guidelines for three kinds of security cooperation assessments: operations assessments (which focus on “are we doing the right things”), task assessments (which focus on “are we doing things right”), and functional assessments (which focus on “are we efficient and effective”).⁹

Sixth, OSD established theater campaign plans (TCPs) for CCMDs and campaign support plans for the military services as a way of applying military planning techniques beyond traditional war plans to incorporate more steady-state activities like security cooperation. These plans allowed for a Secretary of Defense–led analysis and discussion of security cooperation ends, ways, and means that did not exist before. More recently, the CCMDs have begun to develop country plans that promise a clear auditing trail between theater objectives and the means (activities and resources) employed to achieve them that involve U.S. allies and partner nations, given appropriate country objectives and metrics as well as improved country-level data collection.

CCMD Planning and AM&E Processes

Each CCMD has its own planning and AM&E processes for operational and security cooperation missions. EUCOM, PACOM, and AFRICOM illustrate some of the similarities and differences in approaches.

EUCOM relies on annual line of activity (LOA) reports to show progress with respect to TCP objectives. Largely subjective assessments, the reports include a color-coded summary assessment and narrative sections that describe current status, results, hindrances to accomplishment, and strategy and plan considerations. While the office primarily responsible for the LOAs writes the progress reports, they are reviewed by the embassy-based senior defense official/security cooperation officer and approved by the EUCOM country desk officer. The EUCOM J7 assessments division integrates these LOA progress reports with political and military analysis to generate country and thematic assessments. In addition to this level of assessment, the J7 also prepares the comprehensive joint assessment for EUCOM, part of a common process across the CCMDs to assess progress toward OSD planning guidance objectives and intermediate military objectives (IMOs).¹⁰

PACOM evaluates the progress that the command is making with respect to lines of effort (LOEs) spelled out in its TCP. This LOE-based evaluation process is systematic and complex. IMOs are subdivided into a number of more specific effects that represent environmental conditions resulting from an action or actions that contribute to IMO achievement. Effects are evaluated by measures of effects, generally qualitative criteria, and associated quantitative

⁹ Joint Chiefs of Staff, *Joint Publication 3-20*, draft, 2016, p. 53 (provided by sponsor in summer, 2016).

¹⁰ Marquis et al., 2016, pp. 18–23.

metrics called measures of effects indicators. Effects are further broken down into mission-essential tasks, which constitute the basic steps required to achieve effects. These strategic tasks are evaluated by measures of performance, resources required for mission accomplishment, and capability enablers known as measures of performance indicators. At present, tasks are largely self-evaluated by those responsible for executing them. TCP planners generally rely on proxy qualitative measures rather than direct quantitative measures when doing evaluations of IMOs, particularly those that are “too large to measure quantitatively.” PACOM’s approach includes red, yellow, and green “stoplight” objective assessment charts and comments from subject matter experts (SMEs) that the J83 Assessment Divisions synthesize into a narrative.¹¹

Although EUCOM and PACOM have taken important steps toward improving theater planning and AM&E, there are still gaps and areas for improvement. EUCOM has developed customized tools for prioritizing the many objectives in this complex area of operations, as well as a customized system for tracking progress toward objectives. However, the connections between country and theater planning elements are inconsistent, and there is no direct path from security cooperation activities and events to IMOs. In the area of AM&E, processes are nearly all based on self-assessment or the counting of very basic event-level data. For its part, PACOM has improved theater planning and AM&E by reducing the number of IMOs. The 2015 version of its TCP is much easier to understand and utilize for AM&E than the previous plan. In addition, PACOM has a well-developed framework for evaluating IMOs associated with its major LOEs. In spite of these improvements, the connections among theater campaign planning, security cooperation country planning, and AM&E are not as tight as they could be. Also, the way PACOM conducts AM&E is complex and does not extend to the country level.¹²

While still in the early stage of implementation, AFRICOM’s Integrated AFRICOM Theater Synchronization System appears to offer a clear method for linking the elements of its TCP (LOEs and IMOs) with the components of its country cooperation plans (country-level objectives, focus areas, capability packages, projects, and resources). As part of the planning process, effects associated with IMOs at the theater level are explicitly connected to effects associated with country campaign objectives. Based on functions delineated in doctrine and guidance, SFA focus areas are designed to achieve country-level effects. Each focus area has an intent, specific milestones, and a designated office of primary responsibility (OPR). AM&E-related OPR responsibilities include monitoring and reporting progress on milestones, project funding and implementation, and evaluating performance/effectiveness, as well as supporting SFA country assessments.¹³

The Army’s Independent Approach to Security Cooperation AM&E

Without detailed DoD guidance, the military services have largely pursued independent approaches to security cooperation AM&E. Defining the term broadly, the Army views assessment as an aspect of each phase of the operations process: planning, preparation, and execution. In this respect, security cooperation differs very little from wartime or other contingency

¹¹ Marquis et al., 2016, pp. 23–28.

¹² Marquis et al., 2016, pp. 28–29.

¹³ AFRICOM Integrated Theater Synchronization System information briefing provided to the authors by the DSCA in the spring of 2016.

Figure 4.2
Comparing Army Operations and DoD AM&E Constructs

ADRP 5-0 Operations Process	DoD AM&E
Monitoring: “continuous observation of those conditions relevant to the current operation” through the collection of “information about the current situation that can be compared to the forecasted situation described in the commander’s intent and concept of operations.”	Assessments: refers to baseline assessments for leaders and planners that address questions like these: What is the security environment in which these activities will take place? What are the partner’s existing capabilities, and what does the United States want to improve? How well aligned are the partner’s interests and values with those of the United States?
Evaluation: the use of “criteria to judge progress toward desired conditions and determining why the current degree of progress exists.” Such criteria come in the form of MOEs and MOPs. “MOEs help determine if a task is achieving its intended results. MOPs help determine if a task is completed properly.”	Monitoring: means tracking efforts to determine whether inputs (e.g., money and effort) are translating into outputs (e.g., equipment, training, education, and information). These outputs then serve as the basis for tracking progress toward objectives (i.e., outcomes).
Recommendations: Drawing on the evaluation of progress, a commander’s staff discusses the relative merits of possible improvements to the plan and makes recommendations on a course of action to the commander.	Evaluation: examines outcomes and is crucial to understanding what is working and what is not.

RAND RR2165A-4.2

operations. The Army’s continuous assessment process cycle is described in Army Doctrine Reference Publication (ADRP) 5-0, *The Operations Process*.¹⁴ This cycle includes three basic components: monitoring, evaluating, and recommending action for improvement. As shown in Figure 4.2, this parallels, but is not identical to, DoD’s AM&E construct.¹⁵

Army Security Cooperation AM&E Policy

While not deviating substantially from the assessment approach outlined in operations doctrine, Army security cooperation planners at various levels have taken steps in recent years to apply operational assessment principles to security cooperation. In its Army Regulation 11-31, Army Security Cooperation Policy, published in 2012, Army Headquarters directs the department’s security cooperation organizations to carry out the three assessment functions described in ADRP 5-0, to include establishing MOPs and MOEs for security cooperation activities “as a basis for assessing their progress towards specified security cooperation objectives.” In addition, ASCCs and other Army security cooperation providers are told to enter postactivity security cooperation assessment data into either the CCMD’s theater security cooperation management information system or the Army’s equivalent security cooperation database (both of which have been subsequently merged into G-TSCMIS). Furthermore, as part of the development of the Army Campaign Support Plan, the ASCCs are asked to submit “an assessment of progress towards combatant command objectives” to Army headquarters for review.¹⁶

¹⁴ Department of the Army, *ADRP 5-0: The Operations Process*, Washington, D.C.: Army Doctrine Reference Publication Headquarters, May 17, 2012.

¹⁵ See McNerney et al., 2016.

¹⁶ Department of the Army, *Army Regulation 11-31: Army Security Cooperation Policy*, Washington, D.C.: Army Doctrine Reference Publication Headquarters, March 21, 2013c.

First published in 2013, the *Army Security Cooperation Handbook* elaborates on the security cooperation policy outlined in the aforementioned regulation.¹⁷ With respect to assessment, the handbook provides examples of security cooperation measures of effectiveness and performance that relate to security cooperation focus areas spelled out in OSD policy guidance. In addition, it describes the elements of the security cooperation event evaluation mechanism in Army and CCMD security cooperation databases, beginning with the standard questions that security cooperation activity managers are obliged to answer for all events:

- How would you rate the execution of the event?
- How successful was the event in support of the engagement objectives?
- Should the scope or frequency of the event be increased, decreased, or maintained at the current level?

Recognizing the limited utility of information derived from summary questions that lack objective standards for measurement, the handbook recommends that activity managers use the evaluation tab's free text box to expand on their event evaluation by taking into account MOPs, MOEs, resource shortfalls, barriers to mission execution deriving from authorities, policies and necessary permissions, and the risks to security cooperation resulting from lack of resources or other obstacles. For additional evidentiary support, the handbook suggests attaching an AAR to the event record. Finally, the handbook mentions the lack of a standardized DoD system for evaluating progress toward the achievement of security cooperation objectives at the country level. When parent CCMDs have their own systems, the handbook recommends that ASCCs "adopt and support" them. Where such systems are not yet in place, ASCCs should establish their own means of evaluating their LOAs and tasks that support each CCMD security cooperation objective.¹⁸

Army Security Cooperation AM&E Doctrine

Country plan assessments are emphasized in Field Manual 3-22, *Army Support to Security Cooperation*. Consisting of "a comprehensive analysis of the operational environment," country assessments should be performed "in close collaboration with the U.S. country team, the host-nation government, and other multinational partners." Their purpose is to serve as a "baseline" for identifying needs and measuring progress. In carrying out assessments of a partner country, security cooperation planners should pay particular attention to force development in terms of doctrine, organization, training, materiel, leadership and education, personnel, facilities and policy (DOTMLPF-P) capabilities that relate to country objectives.¹⁹ According to the field manual, an initial assessment of the environment and partner capabilities must be updated through continuous observation of current conditions (monitoring) in order to "judge progress toward desired conditions" and determine "why the current degree of progress exists" (evaluation). In addition, the commander and staff "must continually challenge their original framing

¹⁷ Department of the Army, *APD 11-31: Army Security Cooperation Handbook*, Washington, D.C.: Army Doctrine Reference Publication Headquarters, February 6, 2015.

¹⁸ Department of the Army, *APD 11-31: Army Security Cooperation Handbook*, Washington, D.C.: Army Doctrine Reference Publication Headquarters, March 5, 2013b, pp. 55–59.

¹⁹ Department of the Army, *FM 3-22: Army Support to Security Cooperation*, Washington, D.C.: Army Doctrine Reference Publication Headquarters, January 2013a, pp. 3-25 to 3-28.

of the situation” to ensure (1) that security cooperation activities are in sync with stated objectives and (2) objectives are appropriate for achieving the desired end state.²⁰

Army Security Cooperation Planners’ Course

Army Headquarters G-3/5/7 developed the Army Security Cooperation Planners’ Course to help put into place the security cooperation planning and AM&E procedures and principles described in the above regulation, handbook, and field manual. Oriented toward theater Army security cooperation planners, this one-week course attempts through lecture and practical exercises to provide Army officers with no or limited security cooperation experience with the basic knowledge and skills required to formulate and evaluate (ASCC) country plans that support CCMD theater plans.²¹ The course’s assessment block of instruction (which consists of a one-hour lecture and a two-hour practical exercise) introduces students to assessments of performance and effectiveness and shows how they can be used in security cooperation planning and decisionmaking, drawing on Army policy and doctrine and AM&E best practices found in the policy and social science literature.²²

Applying Army AM&E Policy and Doctrine to Security Cooperation Processes

In the past several years, different organizations within the Army have attempted to take nascent policy and doctrine and apply them to the development of AM&E security cooperation processes. Although conceptually promising in some cases, these initiatives have not always been fully or successfully implemented.

Headquarters, Department of the Army’s Partner Interoperability/Capability Assessments

In the mid-2000s, Headquarters, Department of the Army (HQDA) G-3/5/7 proposed a framework for assessing the connection between a partner’s military capability/capacity and interoperability with the United States, existing building partner capacity programs, and U.S. objectives with respect to building partner capacity in target countries. Known as the Army Security Cooperation Continuum, the purpose of this framework was to provide U.S. Army security cooperation planners with a rough assessment of the ground force capacity and interoperability of every significant country in the world. Relying on in-person facilitated discussions with groups of SMEs from the Army staff and the major commands, G-3/5/7 plotted countries in a linear fashion along a vertical interoperability axis and a horizontal capability axis.

²⁰ Department of the Army, 2013a, pp. 3-25 to 3-28.

²¹ Department of the Army, *U.S. Army Security Cooperation Planners’ Course Student Handbook*, Deputy Chief of Staff, G-3/5/7, Strategy, Plans and Policy Directorate, April 11–16, 2016.

²² The assessment portion of the planners’ course draws from the following RAND reports on DoD security cooperation AM&E: Jefferson P. Marquis, Richard E. Darilek, Jasen J. Castillo, Cathryn Quantic Thurston, Anny Wong, Cynthia Huger, Andrea Mejia, Jennifer D. P. Moroney, Brian Nichiporuk, and Brett Steele, *Assessing the Value of U.S. Army International Activities*, Santa Monica, Calif.: RAND Corporation, MG-329-A, 2006; Jennifer D. P. Moroney, Jeff Marquis, Cathryn Quantic Thurston, and Gregory F. Treverton, *A Framework to Assess Programs for Building Partnerships*, Santa Monica, Calif.: RAND Corporation, MG-863-OSD, 2009; Jennifer D. P. Moroney, Joe Hogler, Jefferson P. Marquis, Christopher Paul, John E. Peters, and Beth Grill, *Developing an Assessment Framework for U.S. Air Force Building Partnership Programs*, Santa Monica, Calif.: RAND Corporation, MG-868-AF, 2010; and Jennifer D. P. Moroney, Beth Grill, Joe Hogler, Lianne Kennedy-Boudali, and Christopher Paul, *How Successful Are U.S. Efforts to Build Capacity in Developing Countries? A Framework to Assess the Global Train and Equipment “1206” Program*, Santa Monica, Calif.: RAND Corporation, TR-1121-OSD, 2011.

Structurally, the continuum's capability/capacity axis had its flaws. HQDA's G-3 concept focused mostly on the size of partner contributions to coalition operations (the hardware aspect of capacity) rather than on software components, such as training and doctrine. It also neglected to specify the full context in which their ground forces might be employed when considering a partner's capacity, focusing on its ability to contribute to international operations as opposed to its capacity to counter domestic threats. Nevertheless, if modified and actually put into practice, HQDA's partner capability/interoperability assessment had the potential to provide useful information to security cooperation planners and programmers who lacked domain expertise or an Army perspective on what is needed from partner militaries. Unfortunately, however, the continuum initiative was dropped in the late 2000s, and the Army has yet to establish a systematic process for injecting service partner country assessments into joint and interagency planning and programming.²³

U.S. Army Pacific's Strategic and Tactical Evaluations

The ASCCs have continued to be laboratories for the development of security cooperation planning and AM&E frameworks. To carry out its mission to build relationships and partner capacity, for example, U.S. Army Pacific (USARPAC) conducts security cooperation strategic and tactical level planning and evaluation for unified land operations across the Pacific theater. Strategic evaluation enables USARPAC to evaluate whether its security cooperation activities have helped to accomplish its strategic objectives, update its set of security cooperation activities in light of lessons learned and progress toward its strategic objectives, and provide feedback and guidance to tactical units to execute their security cooperation–related missions. Tactical evaluation enables executing units to determine whether they have fulfilled their stated mission objectives, learned from the event, and passed along that knowledge to other units, generally through AARs.²⁴

USARPAC's Strategic Management System (SMS) is a comprehensive framework that connects specific security cooperation activities with intermediate-level and ultimately campaign-level objectives, and specifies the measures of performance and effectiveness used to monitor progress toward each campaign objective. Despite the framework's appealing design, the effectiveness of SMS evaluations is constrained by shortcomings in execution. Although the process builds in many steps to draw on cross-echelon monitoring and evaluation, evaluation is predominantly driven by senior officers' judgments of partner nation outcomes. In part, this is due to a lack of strategically relevant measures that are associated with a TOC specifying how each planning task can contribute to progress toward achieving campaign objectives.²⁵ In addition, most MOEs and MOPs in SMS capture quantities of outcomes and outputs rather than their quality, assuming that increased engagement will result in increased capacity.²⁶ Finally, SMS very rarely integrates information from tactical evaluations into strategic evaluations. That said, USARPAC AARs are currently not a very useful source of data for strategic evaluations because those who write them do not assess their performance or their perceptions of the effectiveness of the activity against any standard criteria.²⁷

²³ Marquis et al., 2016, pp. 33–35.

²⁴ Watts et al., 2017, pp. 33–58.

²⁵ Watts et al., 2017, pp. 33–39.

²⁶ Watts et al., 2017, p. 48.

²⁷ Watts et al., 2017, pp. 54–58.

Potential Improvements in Army AM&E

Looking ahead, the Army will be fairly well positioned in terms of DoD and service policy and doctrine to conduct effective security cooperation AM&E. Although the CCMDs are the principal audience for the OSD DoDI on security cooperation AM&E, the military services have important roles to play, particularly with respect to two areas: development of initial country assessments and implementation of what OSD calls the “Initial Design Document for significant security cooperation initiatives.” The CCMDs will be dependent on the services—in particular, the ASCCs—for much of the expertise and data required for the completion of these documents.

As previous sections in this chapter indicate, the Army has already done a good deal of work to prepare for this role in terms of publishing security cooperation guidance that integrates service assessment doctrine with security cooperation planning. The Army’s work to date has provided security cooperation planners with an introduction to best practices in AM&E, and facilitated development of AM&E frameworks that inform the joint community about partner ground-force capabilities and demonstrate how service security cooperation activities contribute to CCMD and global objectives.

However, as Figure 4.3 shows, the Army has more work to do to improve and standardize the processes that would allow for the effective implementation of these frameworks.

First, our analysis of OSD, CCMD, and Army approaches to security cooperation AM&E highlighted that, although many AM&E best practices have been identified, incorporating them into military planning and program management requires an **iterative dialogue** among stakeholders, informed by experimentation, learning, and accountability.

Second, while OSD and the CCMDs play a top-down role in providing AM&E guidance, the Army is well positioned to exert leadership using a bottom-up approach to shaping AM&E processes through **innovative and effective implementation**. The Army should have significant leverage to determine how to meet requirements for initial country assessments, project design, activity monitoring, and utilizing measures of performance and effectiveness.

Third, all aspects of AM&E are important, but the ultimate focus should be on **outcomes**. What value do the Army’s security cooperation investments provide? One reason secu-

Figure 4.3
Potential Army AM&E Improvements

Iterate	Army needs iterative dialogue informed by experimentation, learning, and accountability
Lead by doing	Army can shape DoD’s AM&E processes through innovative and effective implementation
Focus on outcomes	Outcomes are less quantifiable and less predictable than output, yet more important
Lead with initial country assessments	Army is well positioned to provide assessments of country’s capabilities
Improve reporting	Army can work with ASCCs to strengthen evaluation reporting
Prioritize	Army can work with OSD and CCMDs to ensure limited resources for AM&E are applied rationally through vigorous prioritization

rity cooperation is so challenging is that outcomes can be influenced but not controlled. In fact, the most important requirements for successful security cooperation lie with the partner nation. Army staff should be held accountable, but less for overall outcomes and more for their ability to help senior leaders make decisions about how to invest security cooperation resources. Understanding high-value security cooperation outcomes and planning for even better outcomes relies on effective AM&E: assessing partners, monitoring and reporting on activities, designing projects, and supporting evaluations.

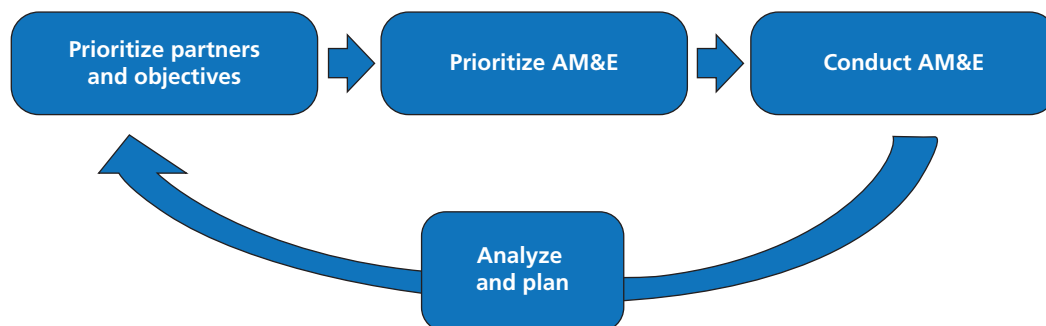
Fourth, the Army should collaborate with OSD, the Joint Staff, CCMDs, and other services on institutionalizing rigorous and widely applicable methods for incorporating service perspectives into **initial assessments of partner capabilities**. The Army could do this by revisiting and building on the example of the Army Security Cooperation Continuum process described earlier. These initial assessments would serve as one of the cornerstones of country planning.

Fifth, the Army could work with these same stakeholders to **improve reporting** on service-executed security cooperation activities so that it is based on strategically relevant measures and tactical evaluations of specific events. In particular, the Army could work with the ASCCs to compare existing AM&E frameworks (e.g., USARPAC's SMS) with the framework developed in this report. Army leaders could then improve the quality and consistency of how ASCCs support evaluations of how well activities advance strategic objectives.

Finally, effective AM&E can be costly and time-consuming. As the Army helps shape DoD-wide AM&E processes, it has an opportunity to ensure that **AM&E resources are applied rationally**, using a “less is more” approach and avoiding using too many metrics for too many partners. The Army should play a central role in DoD's efforts to balance the need for AM&E to be effective with the need to minimize costs and burdens on staff managing programs and executing activities. While AM&E concepts and frameworks would benefit from greater standardization, the type of AM&E applied to activities—especially degree of AM&E intensity—should vary according to service and DoD-wide priorities.

As shown in Figure 4.4, AM&E should both inform security cooperation priorities and be applied based on those priorities. The intensity of monitoring and evaluation, in particular, should depend on the relative importance of partners and objectives, as well as cost of activity and likelihood of success of AM&E. This is a point we return to in Chapters Six and Seven.

Figure 4.4
AM&E and Prioritization



SOURCE: RAND analysis.
RAND RR2165A-4.4

As this section has shown, neither DoD nor the Army in particular is starting with a blank canvas. The intellectual preparation of the battlefield for conducting AM&E is already well under way. The Army's primary challenge, then, is to help shape DoD's AM&E environment and preserve decision space for the Army to implement DoD's broad AM&E framework in the most effective ways possible. Army leaders can do this in several ways, including the following: by providing feedback to other parts of DoD based on real-world experiences, by tailoring DoD-wide guidance to Army needs, and by testing metrics and other components of AM&E through tabletop exercises and other approaches to learning. The analysis and tools provided in this report should help on all these fronts.

Good Practices in Development AM&E

As our review of the security cooperation and international development assistance literatures found, fundamental similarities exist between the objectives and activities of security cooperation and those of development assistance. Therefore, as a first step to developing a framework to help the Army implement an AM&E process in line with good practice and emerging DoD guidance, we reviewed international development assistance organizations' AM&E processes.

This chapter presents lessons learned from the international development literature on AM&E that are relevant for the Army's specific context and needs. We briefly review our research approach and explain why we looked to the international development literature and community of practice for lessons and approaches. We also summarize the trajectory that AM&E has followed in the broader development community, and how the Army can benefit from the successes and failures the development community experienced along the way.

The remainder of the chapter discusses how the development community approaches AM&E and what constitutes high-quality AM&E. This includes a detailed summary of components that development institutions use to carry out AM&E. We conclude with a set of recommendations for how the Army can leverage these AM&E tools and approaches.¹

AM&E in the development assistance community faces needs and challenges similar to those of AM&E in the Army. Development agencies have internal and external requirements to measure results on a regular basis to inform operations, prioritize programs, and support accountability. These organizations use different metrics, often lack unified assessment systems, and grapple with how to tailor programs, objectives, and outcomes to meet the needs of foreign partners. Much like ASCC planners, development organizations are working to set clear achievable metrics, develop assessment frameworks that are calibrated to program needs and constraints, and increase cooperation with partner institutions, especially in host countries.²

At the same time, the international development community has goals and constraints that are distinct from those of the Army when it comes to AM&E. Development institutions face different political constraints and generally undertake activities that are relatively less sensitive than security cooperation operations. An important difference between how the development community designs its activities and Army security cooperation is the role of generating public knowledge. Many development institutions, including the World Bank and

¹ In contrast to Chapter Two, here we take a broad view of the development literature, beyond cross-country aid effectiveness. We focus more on tools to support microlevel assessment, given those approaches are more relevant to designing an activity-specific AM&E for security cooperation.

² Elliot D. Stern, Laura Altinger, Osvaldo Feinstein, Marta Marnon, Daniela Ruegenberg, Nils-Sjard Schultz, and Nicolai Steen Nielsen, "Thematic Study on the Paris Declaration, Aid Effectiveness and Development Effectiveness," *Evaluation of the Paris Declaration*, November 2008; OECD, *The Paris Declaration on Aid Effectiveness and the Accra Agenda for Action*, 2005/2008.

the Department for International Development in the United Kingdom, have as part of their mission expanding what the public knows about foreign aid effectiveness and even AM&E methods.³ They invest resources in how to improve their own programs and increase resource efficiency, while also spending resources to share information and learning. The Army places less emphasis on investing in these types of public goods.

Despite the differences in programming between the Army and the international development community, we believe there are relevant lessons learned and good practices from the development community's sizable investments in AM&E. Therefore, we reviewed the trajectory of AM&E and current approaches and lessons learned for the international development community. In particular, we conducted desk reviews of and engaged international development experts at RAND, the Inter-American Development Bank (IDB), the World Bank Group (WBG), the Millennium Challenge Corporation (MCC), the United States Agency for International Development (USAID), and DoS. We collected and reviewed guidance documents and frameworks used by these institutions, and we drew on the broader international development literature to understand the trajectory of AM&E use and their best practices and challenges with AM&E.

AM&E Trajectory in International Development

After the Cold War, new debates emerged about foreign aid effectiveness, the need for increased accountability, and the desire to rethink donors' approaches to aid allocations.⁴ The first Organization for Economic Cooperation and Development (OECD) countries to grow their evaluation culture and AM&E systems were fundamental to spreading knowledge, information, and training, and for launching evaluation institutions.⁵ The OECD countries that followed these trends were incentivized by the first movers and by international organizations.⁶ As a result, development agencies have faced a larger demand for evidence-based policies, effectiveness, results measurement, and causal impact estimates from donors, program beneficiaries, and oversight organizations.⁷ AM&E systems have been strengthened, and even though progress has not been homogeneous or linear,⁸ AM&E systems are now being implemented more systematically and strategically in development organizations around the world. As a result of collective action for better data collection and analysis, methods for AM&E have improved and many development agencies are now not only addressing TOCs and activity performance, but also addressing whether those activities causally affect outcomes and impacts.⁹

³ Development Impact Evaluation, *Transforming Development Through Impact Evaluation, Annual Report, March 2015 to March 2016*, World Bank, 2016; Department for International Development, "DFID Research Strategy 2008–2013," n.d.

⁴ Stern et al., 2008.

⁵ Jody Zall Kusek and Ray C. Rist, *Ten Steps to a Results-Based Monitoring and Evaluation System: A Handbook for Development Practitioners*, Washington, D.C.: World Bank Publications, 2004.

⁶ Kusek and Rist, 2004.

⁷ KPMG International, *Monitoring and Evaluation in the Development Sector, A KPMG International Development Assistance Services (IDAS) Practice Survey*, 2014; INTRAC, *Monitoring and Evaluation: New Developments and Challenges*, International Conference, Soesterberg, The Netherlands, June 14–16, 2011.

⁸ Kusek and Rist, 2004.

⁹ KPMG International, 2014.

How Development Organizations Implement AM&E

A range of U.S. and non-U.S.-based organizations are using robust AM&E systems, although those systems vary by organization. U.S. development organizations, such as USAID, MCC, and the State Department, are incorporating AM&E into their activities and increasingly promoting their use. To better understand the current use of AM&E in development agencies, we spoke with WBG, IDB, USAID, and State Department staff to learn about their AM&E frameworks. In addition, we researched practices within these organizations and MCC. Overall, the emerging DoD AM&E guidance discussed in Chapter Four follows the same principles as those underlying AM&E approaches at WBG, IDB, USAID, and the State Department.

Multilateral Organizations

WBG has developed an AM&E system that it uses for accountability, performance management, and learning.¹⁰ This system was established for monitoring and evaluation purposes at four operational levels: activity, program, country, and corporate. WBG staff who implement activities conduct “self-evaluations” of their work systematically, and these self-evaluation systems operate well enough that they allow learning and timely adaptation throughout the activity life cycles. Besides the self-evaluations conducted by staff, WBG also has an autonomous entity, the Independent Evaluation Group, which reports directly to WBG’s board of directors. The role of the Independent Evaluation Group is to ensure that independent and objective evaluations are being conducted on the work of WBG, and to identify and disseminate lessons learned from experience.¹¹

IDB is another organization that has well-developed AM&E systems that are integrated into its operations and activity life cycles. IDB has a monitoring and evaluation structure that is quite similar to WBG’s. IDB management focuses on performance evaluations and monitoring, while an autonomous entity that reports directly to the board of executive directors, the Office of Evaluation and Oversight, monitors implementation and evaluates IDB’s work after activity completion.

Both WBG and IDB have developed AM&E systems that make it possible to incorporate their AM&E framework into their activities and programs from the outset. These systems work in cycles: at activity inception, evaluation specialists or other staff members ensure that activity indicators have baselines and targets—goals or expected outcomes. For particular targets that are difficult to estimate, staff may also use existing predictive models or evidence to come up with those estimates; at IDB, activities also undergo cost-benefit analysis before the activities start.

Setting up these baselines and targets allows for activity monitoring and supervision throughout the entire activity duration (i.e., calculating the share of targets that have been met), activity course correction in case it is needed, and performance evaluations once activities have finalized. Activities are graded at completion to ensure that targets are met and to compare how actuals compare with targets. Reports are built from these evaluations, and these reports are shared for learning purposes and to inform activity and program donors on the organization’s operations.

¹⁰ Independent Evaluation Group, “Approach Paper Report on Self-Evaluation Systems (ROSES 2016),” *World Bank Group*, No. 95189. February 24, 2015.

¹¹ Independent Evaluation Group, 2015.

U.S. Development Organizations¹²

MCC has developed an AM&E framework for all stages of the activity life cycle, where results and learning are used as a feedback loop that improves future activities and supports accountability.

MCC's distinctive model involves close collaboration with partner nations throughout all stages of the AM&E framework. The AM&E cycle starts by identifying the key constraints that hinder economic growth in the partner nation, and then selecting specific activities that might foster economic growth. A beneficiary (partner nation) analysis is also conducted and ex ante activity benefits and costs are estimated through an economic rate of return analysis. Implementation relies principally on the implementing partner nation agency.

MCC monitors activities through established tools/templates that help staff identify targets, and it tracks performance measures through an ongoing monitoring process. MCC also conducts performance and impact evaluations, and the decision on which type of evaluation to choose is based on activity characteristics: learning potential, feasibility, level of stakeholder commitment, timing of evaluation and activity, and level of coordination between MCC planners and the partner nation.

Independent, external evaluation experts usually conduct evaluations with the aim of ensuring objectivity while answering questions about effectiveness.

USAID has a commitment to accountability and learning for effectiveness and therefore implements “rigorous and quality program evaluation—the systematic collection and analysis of information to improve effectiveness and inform decisions about current and future programming.”

USAID's evaluations are planned in conjunction with activities, and they are implemented by third-party organizations. USAID determines which activities will require an evaluation by analyzing the size of the activity, whether the activity is innovative, whether the activity is a pilot that is planned for expansion, and the importance of the question that will be answered through evaluation. USAID has been able to implement this framework by working closely with its leadership, improving staff capacity in AM&E methods, engaging with outside AM&E networks and communities of practice, and integrating evaluations into their activity life cycles.¹³

The Use of AM&E in Practice

The remainder of this chapter summarizes some of the core methods used to conduct AM&E by the international development community and assesses their relevance to Army security cooperation. These best practices include TOCs, monitoring, performance evaluation, and impact evaluation. We first describe lessons learned from implementing AM&E in the development community. After that, we review each method's function, the resources needed to implement it, and common misperceptions, especially those relevant to the Army.

¹² USAID content based on USAID, “USAID Evaluation,” 2016b; USAID, *Strengthening Evidence-Based Development, Five Years of Better Evaluation Practice at USAID 2011–2016*, 2016a. MCC content based on Millennium Challenge Corporation, *Our Impact*, 2016.

¹³ USAID, *Evaluation, Learning from Experience: USAID Evaluation Policy*, January 2011.

Lessons Learned from AM&E in the Development Community

While we discuss lessons learned that have emerged in the development community in greater depth in Chapter Two, below is a list of lessons that could be useful to Army security cooperation M&E:

- **M&E should not be “one size fits all.”** AM&E activities have different strengths and serve different functions; AM&E planning should acknowledge and leverage these differences.
- **“Return on investment” in learning should drive AM&E methods.** AM&E supports learning, and learning has both benefits and costs. Where there is high value to learning more about an activity or set of activities, the associated investment in AM&E should be higher.
- **High-quality theories of change are critical.** Before conducting AM&E, implementers need to map out an assistance activity’s design. A TOC is a tool to make the design explicit, identify critical assumptions, and guide AM&E plans.
- **There is growing recognition of the value of impact evaluations.** Impact evaluation, an AM&E tool focused on establishing causal effects, helps an organization understand whether and how an assistance activity works, using high-quality evidence. Impact evaluations are not needed or appropriate in all situations, but they are increasingly a core tool for a broad range of organizations.
- **All activity planning should incorporate AM&E.** Every activity needs some form of AM&E, even if it is limited to monitoring. AM&E needs to be consistently integrated with activity implementation, especially at the planning stages.
- **Qualitative and quantitative evaluation are useful on their own and even more beneficial when used in combination.** Quantitative and qualitative methods can answer questions about program effectiveness, and when used in conjunction they allow triangulation of results and a deeper understanding of findings.
- **Strong collaboration with partner nations is crucial for effective AM&E.** The Army may work with partners to reach mutual understanding of the level of compromise that is expected from AM&E activities, and agree on sharing its costs and benefits before undertaking arduous work together.

AM&E Approaches in the International Development Community

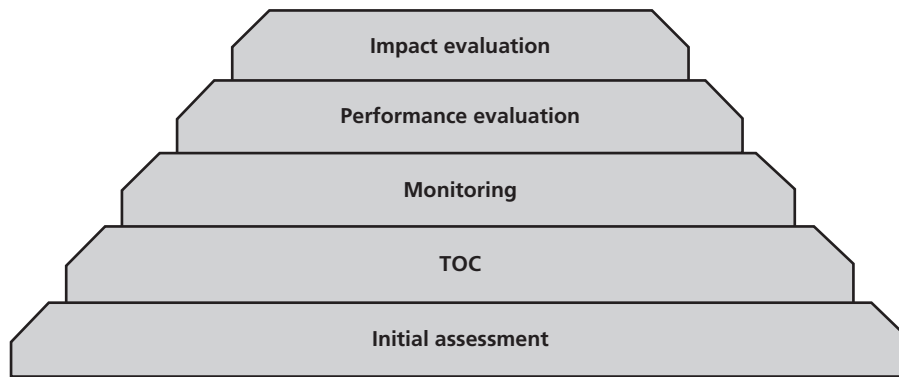
Figure 5.1 shows the approaches covered in this section, and the order in which they are covered: TOC, monitoring, performance evaluation, and impact evaluation.

Initial Assessment

An initial assessment may be conducted before establishing a TOC. An initial assessment helps justify the existence of an activity or program. It helps with identifying a gap between *what is* and *what is desired* in terms of capabilities and outcomes. After the gap is identified, potential causes for the gap and strategies to fill the gap are identified.¹⁴ Planners need to select the best alternative among those strategies based on their expected benefits, risks, and indirect effects. Once the best alternative has been identified, the assessment will provide information

¹⁴ Ryan Watkins and Maurya West Meiers, *A Guide to Assessing Needs: Essential Tools for Collecting Information, Making Decisions, and Achieving Development Results*, New York: World Bank Publications, 2012; Office of Migrant Education, “Comprehensive Needs Assessment,” New Directors Orientation, 2001.

Figure 5.1
Pyramid for AM&E



SOURCE: RAND analysis and Paul, 2013.

RAND RR2165A-5.1

on the partner nation's ability to implement the alternative, the risks, contextual factors, and any other information that may nurture or affect implementation and that might be useful for monitoring and evaluation purposes. The development community usually refers to the initial assessment as "needs assessment" because it identifies a need—or gap—and the best solution to be implemented.

The development community identifies the initial assessment as part of a cycle, in which the assessment informs the TOC, which is followed by planning, implementation, monitoring, and evaluation.¹⁵ Evaluations then provide valuable feedback about activities that have been implemented, and these insights are used to inform future TOCs and their initial assessments.

The literature sometimes identifies initial assessments as part of the TOC approach, and other times as an independent step of the AM&E life cycle. Figure 5.2 shows how the initial assessment looks in our analysis, which is when it supports the TOC rather than being an independent step.

Resource Intensiveness

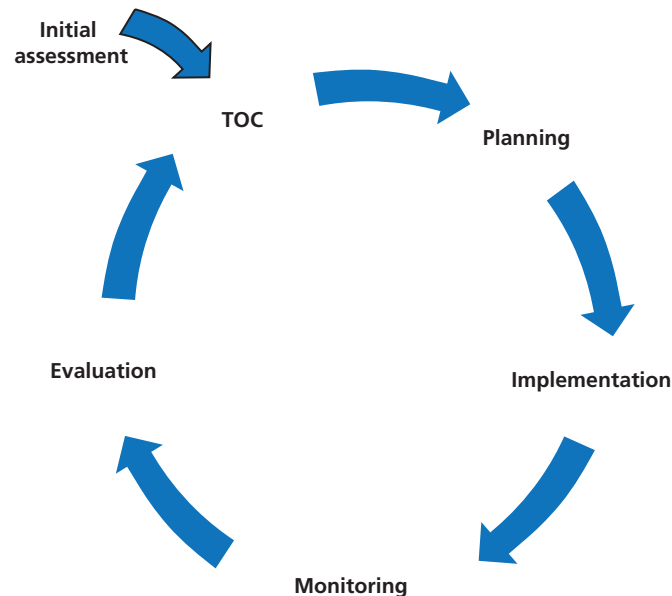
An initial assessment is critical to designing and implementing effective security cooperation that meets the planner's objective, but it is not necessarily costly to implement. The assessment can be as simple as talking with people knowledgeable about the partner nation and its needs, a process that can be done at almost zero cost. In contrast, some situations may benefit from extensive initial assessments that involve deeper research or even site visits. In general, the higher priority the partner nation is and the less security cooperation the Army has done with that country, the more the planner will benefit from a robust initial assessment of conditions and needs.

Misperceptions

Despite the name, an initial assessment is not a one-off activity. Where a planner is working with a new country or a country that the Army has not engaged with recently on security cooperation, a thorough assessment sets the stage for planning security cooperation activities. However, periodic follow-up "initial assessments" should be done when goals evolve or change, the security cooperation planner is seeking to implement activities not done before, or the results of a previous initial assessment are dated. Additionally, an initial assessment should not be done

¹⁵ Watkins and Meiers, 2012.

Figure 5.2
AM&E Cycle



SOURCE: Authors' elaboration based on Watkins and Meiers, 2012.
RAND RR2165A-5.2

and put on a shelf. Once security cooperation activities based on the assessment have been implemented, the planner should revisit the initial assessment to verify (or refute) assumptions and confirm whether identified needs and gaps were accurate.

TOC

Development organizations increasingly require a TOC for programs and activities they carry out, and organizations that support program evaluation similarly value and demand TOC-based evaluation designs. A TOC is a conceptual tool that maps out how an activity accomplishes the implementer's desired goals in a particular context.¹⁶ TOCs are the basis for other evaluation activities, and that is why they are at the bottom of the pyramid in Figure 5.1. They are the basis for verifying whether an organization is accomplishing expected milestones and for developing lessons learned after activity implementation; they also have accountability and transparency purposes. They should be developed during the planning stage of an activity, while there is still time to make adjustments to the design and the AM&E approach. We discuss the TOC in the context of an activity, but the approach applies equally well to groups of activities and higher-level plans, like tasks. When conducted carefully and thoroughly, a TOC provides multiple benefits to the implementer and evaluator:

- It requires the underlying “logic” for the activity to be explicitly and clearly written down, which often makes gaps or limitations in the design clear.¹⁷

¹⁶ Center for Theory of Change, “What Is a Theory of Change,” n.d.

¹⁷ A TOC is similar to a logic model. The terms are sometimes used to refer to related but distinct concepts. Here we do not discuss logic models but instead use *TOC* to refer to the broad class of tools and methods. Dawn Roberts and Nidhi Kattri, *Designing a Results Framework for Achieving Results: A How-To Guide*, Independent Evaluation Group, Washington, D.C.: World Bank, 2012.

- It makes explicit the assumptions surrounding the activity, including actions expected of other stakeholders, resource requirements, and complementarities with other activities.
- It forces the implementer to think about potential unintended consequences and provides an opportunity to identify ways to mitigate those effects when they are undesirable.
- It provides the foundation for an AM&E approach. The TOC highlights where an activity relies on strong assumptions or where there is weak evidence for a relationship between inputs, outputs, and outcomes. An AM&E approach for that activity would seek to fill those knowledge gaps.

A TOC has multiple sequential components, and it is typically constructed starting with the desired goals and mapping backward to identify the outcomes, outputs, and activities that must occur to achieve the goals.¹⁸ In Chapter Six we review in more detail the process for constructing activity-level TOCs, including examples for security cooperation use.

The initial assessment can also be conducted at different levels (e.g., country level or activity level) because it depends on the level of the activity or program being implemented. In Chapter Six we describe how to apply initial assessments to Army security cooperation.

Resource Intensiveness

TOCs are not costless, but their benefits tend to outweigh their costs.¹⁹ The cost of creating a TOC includes up-front training costs and the time spent developing the TOC for each activity, but these costs are relatively modest compared with the long-term costs of conducting ineffective activities, poorly implemented activities, or low-value AM&E. If created during the planning stages of an activity, a TOC can help avoid costs during later implementation stages by flagging the necessary inputs and actions, including actions by others that are on the activity's critical path. The TOC will also help ensure AM&E resources are used effectively and in ways that maximize learning. In some cases, creating a TOC can identify that an activity is not well designed or not appropriate for the proposed use, saving resources from changing, delaying, or canceling the activity. Finally, TOC costs scale with activity complexity; simple activities will require simpler TOCs, using fewer resources.

Misperceptions

The TOC is often perceived as too complicated to use or not relevant to the specific program being implemented. A variation on this view is that it is not worth constructing a TOC for simple activities. We disagree strongly. This view undervalues the process of mapping out a TOC, and it ignores the fact that a simple activity typically requires only a relatively simple TOC. If the TOC for an activity is perceived as “too complicated,” then this should be a signal for planners that the activity under consideration is also complicated—and is more, rather than less, in need of a clear statement of how the activity is expected to accomplish the implementer's desired goals. Moreover, once a TOC has been constructed for an activity, then when that same activity is implemented in the future, it may require only minor modifications to the TOC.

¹⁸ Center for Theory of Change, n.d.; Jim Parsons, Caitlin Gokey, and Monica Thornton, *Indicators of Inputs, Activities, Outputs, Outcomes and Impacts in Security and Justice Programming*, UK Aid, Department for International Development, October 15, 2013.

¹⁹ Craig Valters, “Theories of Change: Time for a Radical Approach to Learning in Development,” Overseas Development Institute, September 2015; Duncan Green, “What Use Is a Theory of Change? 6 Benefits, and Some Things to Avoid,” *People, Spaces and Deliberation*, World Bank, 2013; Center for Theory of Change, n.d.

How TOC Fits with Army Current Practices

Army security cooperation would benefit from broad-based use of TOCs at the planning stage. TOCs can be integrated into security cooperation planning and become an essential tool, both at the country and activity levels. ASCC planners would need to be trained on methods to develop a TOC, and they should allocate time to develop TOCs during the planning stage. In Chapter Six, we present a template for an Army-specific security cooperation TOC and examples of how to use it for security cooperation activities.

At the country level, TOCs would take the ASCC country assessments as inputs and build a TOC for each GCC-determined IMO. This high-level TOC flows down and guides the planning around individual or groups of security cooperation activities. There could be a TOC for each security cooperation activity, and these TOCs may be adapted across similar or repeated activities to increase efficiency. Finally, TOCs may be used to inform other AM&E processes and tools.

Monitoring

Monitoring is the systematic and typically recurring collection of information that is conducted while an activity is being implemented. Over time, development agencies have faced an increase in the demand for real-time activity feedback, and this has also boosted the need to conduct continuous monitoring.²⁰ Agencies require TOCs to already be in place before they can start monitoring (see Figure 5.1) because TOCs define the inputs, activities, outputs (and corresponding MOPs), and outcomes (and corresponding MOEs) that will be tracked and measured through continuous monitoring.

Monitoring is valuable because the real-time data it generates can be used to improve activity design and implementation, while activities are carried out, and support other types of evaluation.²¹ Monitoring adds value by gathering continuous information while activities take place, both quantitative and qualitative, about “how much” of an activity was done, where the activity was done, and who conducted the activity.²² Typically, monitoring is the practice of collecting this basic information, but it can also include more complex or sophisticated data collection, focusing on how an assistance activity was received or taken up by recipients, what external factors acted on the activity, or what unintended consequences derived from the activity. Monitoring can also be used to gather data about how an activity is administered or on the mechanisms through which an activity aims to achieve the desired goals. For this reason, monitoring is particularly valuable for tracking progress and informing day-to-day management decisions.

Measures used to monitor an activity should follow the SMART criteria.²³ These criteria describe that each indicator should be Specific, targeting only the outcome it was designed to measure; Measurable, so that it can be quantified or estimated; Achievable, such that it is feasible and attainable; Relevant and Results-oriented, for it to contribute to the overall strategic goals; and Time-bound, which means that there is a deadline for completion of each indicator.²⁴

²⁰ United Nations Development Programme, “Innovations in Monitoring & Evaluating Results, Discussion Paper,” November 5, 2013.

²¹ Independent Evaluation Group, 2015; MIT, “Monitoring and Evaluation: What Can They Do for Me?” n.d.

²² Martha R. Burt, Adele V. Harrell, Lisa C. Newmark, Landon Y. Aron, and Lisa K. Jacobs, “Chapter 6: Choosing an Evaluation Design,” in *Evaluation Guidebook*, Washington, D.C.: Urban Institute, 1997.

²³ McNerney et al., 2016; Operational Policy and Quality Department, *Results Framework and Me&E Guidance Note*, Washington, D.C.: World Bank, April 9, 2013.

²⁴ McNerney et al., 2016; Operational Policy and Quality Department, 2013.

Monitoring is valuable, but it is not a substitute for other evaluation activities. Monitoring informs activity design and implementation while the activity is still ongoing, while evaluation informs future activities.²⁵ In addition, data from monitoring activities often need additional information or context before they can be used to make judgments about effectiveness. For example, collecting information on the number of activities administered or the cost of those activities is valuable, but that information alone does not allow the implementer to draw conclusions about whether the activities achieved the desired long-term objectives.

Most importantly, monitoring conducted outside of a well-designed evaluation cannot be used to make statements about whether an activity worked. For example, knowing that ten people participated in a certain training activity does not allow the implementer to say that the training “worked”; it only provides evidence that an activity was carried out.

Monitoring data can come from a range of sources. A common way to collect monitoring data is through administrative sources, through basic records kept on activities as they are implemented.²⁶ Administrative data often focuses on inputs—how much money was spent or training materials used—and can be relatively easily expanded to include data on activities (e.g., trainings conducted), outputs (e.g., number of individuals who complete a training), and sometimes outcomes (e.g., individuals are actually applying the training content in practice). For more sophisticated monitoring, such as outcomes or unintended consequences, the implementer or evaluator may need to undertake dedicated, specific data collection using interviews or surveys.

As long as data collection methods are consistent, monitoring can be used to make comparisons across activities or within activities against targets.²⁷ For example, a survey that asks training recipients to rate their experience can be used over time and across similar types of training to gauge satisfaction or identify context-specific problems, like language barriers. Similarly, if the implementer sets targets for the activity, monitoring can be designed and carried out to determine whether targets were met. In the case of a training program, monitoring that provides consistent data on trainee performance can track outcomes over time and ensure performance targets are met across different trainers.

Therefore, monitoring can provide substantial benefits to the implementer and the evaluator, and some form of monitoring is critical for all activities. Monitoring enables oversight, supports transparency, and promotes awareness within and across activities; it can help stakeholders make better decisions as activities are carried out, and it can inform how activities are designed and implemented in the future. Not least, monitoring underpins many types of evaluation, providing the data with which decisions about effectiveness and impact get made. In particular, monitoring can provide answers to these questions:²⁸

- How is the activity operating (e.g., is it efficient)?
- Are designated targets being met?
- Are expected outputs and outcomes likely to be produced?

²⁵ MIT, n.d.

²⁶ Mary Kay Gugerty, Dean Karlan, and Delia Welsh, *Using Administrative Data for Monitoring and Evaluation*, Goldilocks Deep Dive, Innovations for Poverty Action, 2016.

²⁷ Paul J. Gertler, Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeesch, *Impact Evaluation in Practice*, Washington, D.C.: Inter-American World Bank Publications, 2011.

²⁸ Burt et al., 1997.

- Are MOPs showing the activity is being completed as expected?
- Who is implementing or delivering the different activity components?
- Who benefits from the activity?
- To what extent are activity beneficiaries satisfied with the services received?

Resource Intensiveness

Monitoring can be costly, and the cost depends on the time it takes to collect data, especially in the case of specialized data collection, and on the required human resources and other data collection infrastructure. Costs also depend on the size of the activity being monitored, the number of indicators collected, and the length of the activity, among other factors. The costs typically fall into three groups: design of monitoring activities, implementation of monitoring, and management or interpretation of data once they are collected. The implementation of monitoring and management of data can require specific technical expertise that might increase costs, similar to the use of sophisticated data collection methods, such as in-depth surveys or purchasing satellite imagery.

Nevertheless, because monitoring can be scaled, it is usually possible to undertake monitoring in a way that ensures costs are less than the benefits. The main risk to the cost-benefit calculation for monitoring is that costs are often well defined for a given level of monitoring, while benefits depend entirely on whether and how the evaluator uses the data that get collected.

Misperceptions

The main misperceptions with monitoring are (1) that it is always costly to conduct and (2) that monitoring can establish *causality* (i.e., whether the outcomes are *caused* by the activities that took place rather than something else). Monitoring can be costly, but the sophistication—and thus cost—can be adjusted based on how complex the activity is and what the benefits of learning are. A simple activity that needs only basic oversight can probably be monitored solely through objective after-action reporting. Where monitoring is costly, that cost should be justified through the benefits of using the gathered information about the activity.

Monitoring cannot and does not establish *causality*, although it can support evaluations that determine the causal effect of an activity on the outcomes of interest. In other words, measuring what takes place when a security cooperation activity gets implemented or collecting data on the outputs of an activity should not be interpreted as evidence that the activity is achieving its outcomes or supporting IMO. Monitoring must be combined with a high-quality evaluation strategy to provide insight into whether an activity works to achieve its goals. In Chapter Six we provide more guidance on specific activities that require some type of monitoring.

How Monitoring Fits with Army Current Practices

Monitoring should have a central role in security cooperation planning and implementation. It is important for monitoring to be fully integrated into security cooperation planning and that all security cooperation activities are monitored, even though the type of monitoring will vary across activities. Some activities require basic monitoring through AARs, while others require more intensive and systematic monitoring. It is desirable to set monitoring standards across activities at the planning stage because this allows consistency across monitoring indicators and processes and comparability for the feedback generated.

Once implemented, a robust monitoring system would generate real-time information that is critical for security cooperation reporting and auditing purposes, and it would provide

key information for evaluations. This monitoring system would track how inputs are used, how activities take place, and whether MOPs—and sometimes MOEs—are being completed as expected. More information on the level of monitoring that is required and how to conduct monitoring for security cooperation activities will be covered in Chapter Six.

Performance Evaluation

Performance evaluation is a tool that helps establish whether activities are implemented according to the activity-level TOC and whether activity implementation, process, and delivery are conducted as planned.²⁹ Concerns about accountability and the desire to use public funds effectively have made performance evaluation a critical approach to increase the quality of development aid.³⁰ Performance evaluation is different from monitoring, which focuses on collecting information on an ongoing basis, because it measures and assesses activity information at discrete points in time, often at the middle or end of an activity.³¹ It is the third approach included in Figure 5.1, and performance evaluation draws on the information about inputs, activities, outputs, and outcomes from monitoring efforts.

Resource Intensiveness

The cost and duration of performance evaluations vary, but they usually increase with activity size, number of people involved, geographic remoteness of the activity location, and activity duration. Performance evaluations can be implemented faster and at relatively lower costs when they rely on existing administrative data that have been collected as part of an ongoing monitoring effort, but organizations implementing performance evaluations may still need to collect additional data that complement earlier monitoring efforts. Some types of performance evaluations can be conducted using relatively simple methods or qualitative data collection, which can reduce cost. For example, an evaluator can shift from surveys to direct observation or focus groups, and self-administered surveys can be used instead of having a team of people administer the questionnaires. These shifts should not be done mid-evaluation, but in the design phase an evaluator has substantial flexibility to balance evaluation needs and costs. The amount of data collected, and the frequency with which the data are collected, can also be reduced.³²

Misperceptions

A common misperception about performance evaluation is that it is sufficient to establish whether activities have an impact on—have *caused*—outcomes. Performance evaluation measures outcomes, but its methods of data collection and analysis cannot establish whether the outcomes are caused by the activities that took place or by something else. The value in performance evaluation comes from what it can reveal about activity design, execution, and delivery that allows timely activity course correction,³³ accountability, and long-term improvement.³⁴

²⁹ J-PAL, “Introduction to Evaluations,” Impact Evaluation, 2016.

³⁰ OECD-Development Assistance Committee, *Principles for Evaluation of Development Assistance*, Paris: OECD, 1991.

³¹ Operational Policy and Quality Department, 2013.

³² Michael Bamberger, Jim Rugh, and Linda S. Mabry, *Real World Evaluation: Working Under Budget, Time, Data, and Political Constraints*, Thousand Oaks, Calif.: SAGE Publications, 2013.

³³ J-PAL, 2016; World Bank, *Monitoring and Evaluation: Some Tools, Methods & Approaches*, Operations Evaluation Department, Washington, D.C.: World Bank, 2004.

³⁴ David Bonbright, Fred Carden, Sarah Earl, Sanjeev Khagram, Nancy MacPherson, Zenda Ofir, and Patricia Rogers, *Impact Evaluation for Improving Development (ie4id)*, 2010.

Performance evaluation is often confused with monitoring, but they are complements rather than substitutes. While monitoring is a continuous process that allows real-time feedback and provides valuable information for evaluation, evaluations are implemented at discrete points in time and may also inform future activities.

How Performance Evaluation Fits with Army Current Practices

A performance evaluation can be helpful for Army security cooperation since it can clarify whether the activity is being implemented according to plan—consistent with the activity-level TOC. Performance evaluation also provides valuable information to improve activity design and implementation; it allows midcourse correction and provides useful feedback to improve future activities. Not all activities require performance evaluations, and activity characteristics such as length of time, resources involved, recurrence, and strategic importance will help determine whether performance evaluation is required.

To implement performance evaluations, the Army should consider following specific guidelines that have been widely applied by OECD countries' development aid:³⁵

1. Establish a clear evaluation policy that comprises the evaluation methods to be used and the role that evaluation plays within the institution.
2. Conduct evaluations impartially and independently from the interests of those executing or managing activities.
3. Make evaluation results available to the public, and ensure recommendations are used and applied.
4. Collaborate closely with partner nations in conducting evaluations.
5. Plan for performance evaluations at activity inception, so that evaluation planning, data collection, and analysis are implemented in tandem with the activity.
6. Incorporate performance indicators from evaluations into monitoring documents and processes. Monitoring and evaluation should be linked and support each other.

We acknowledge that points 2 and 4 may seem conflicting, that working closely with partner nations jeopardizes independence and impartiality. This is an issue that the development community is still working through, but the current state of the art is to have an independent evaluator that is free from financial conflicts with the implementer, which may be the partner nation. At the same time, the evaluator ideally participates in the design process for the activity, which helps ensure the evaluation is effective. This approach requires balance, since there is inherent tension between independence and coordination between the two parties.

Impact Evaluation

The last component of the AM&E pyramid in Figure 5.1 is impact evaluation. Impact evaluations have seen exponential growth since the end of the 1980s,³⁶ mostly due to the need to understand and measure the impact of aid activities on specific outcomes, such as reductions in violence, improvements in health, and reductions in poverty.

Impact evaluation focuses explicitly on identifying the causal impact of an activity on outcomes of interest, but this process—often referred to as *estimating impact*—is not trivial.

³⁵ OECD-Development Assistance Committee, 1991.

³⁶ Drew Cameron and Jorge Miranda, "Trends in Impact Evaluation: Did We Ever Learn?" *International Initiative for Impact Evaluation*, 2015.

Estimating impact cannot be accomplished solely by conducting monitoring or conducting a performance evaluation. Determining impact requires the use of more rigorous impact evaluation methods, which are now well established in the development community. Impact evaluation determines the changes in outcomes that are caused solely by the activity in question and not by something else.

The key to impact evaluation—and to establishing causality—is measuring activity effectiveness using methods that compare outcomes of activity participants with the outcomes participants would have had if they had not participated in the activity. However, observing the same individuals as participants and nonparticipants is impossible. Impact evaluation solves this problem by using methods to construct a group that closely resembles what would have happened to activity participants—the treatment group—if they had not participated. This *comparison* group is also known as a control group. By constructing this control group that closely resembles the group of participants, impact evaluation ensures that the only difference—in statistical terms—between the characteristics of participants and those of nonparticipants is participation in the activity, and any differences in outcomes between both groups can therefore be attributed solely to the activity. In addition, impact evaluation makes it possible to estimate the magnitude of the impact and to identify how the activity works to accomplish impacts. Finally, an impact evaluation provides information about the effectiveness of an activity, which can be useful to compute the cost-effectiveness (i.e., activity effectiveness per dollar invested); when impact evaluations are performed across multiple activities, they would allow comparability across activities in terms of effectiveness per dollar invested.

How accurately the evaluation is able to estimate an activity's impact depends on how accurately the control group resembles the characteristics of participants. The challenge for impact evaluation is building the control group, and it requires using specific methods.

There are multiple ways to implement an impact evaluation, and the best approach will depend on the activity design and the evaluation requirements and constraints.³⁷ One well-known approach, popularized in social sciences in the past 45 years, is the randomized evaluation.³⁸ Development researchers have taken randomized evaluations and applied them to a range of activities and programs around the world.³⁹ A randomized evaluation randomly allocates some individuals into treatment (e.g., some individuals are offered treatment, are encouraged to participate in treatment, or participate in treatment by default) and some into control groups (the control group closely resembles what would have happened to activity participants if they had not participated), which ensures that on average these two groups (treatment participants and control groups) have similar characteristics and the only difference between them—in statistical terms—is activity participation.

While there are important advantages to randomized evaluations, they are not always feasible to implement.⁴⁰ Impact evaluation can be conducted using a range of alternative methods

³⁷ Kene Ezemenari, Anders Rudqvist, and K. Subbarao, *Impact Evaluation: A Note on Concepts and Methods*, World Bank Poverty Reduction and Economic Management Network, Washington, D.C.: World Bank, 1999.

³⁸ Judith Gueron, "The Politics and Practice of Social Experiments: Seeds of a Revolution," in *Handbook of Field Experiments*, Amsterdam: North Holland Publishing Co., March 21, 2016.

³⁹ Luc Behaghel and Philippe Zamora, "Randomized Controlled Trial: a Historical Perspective," *Abdul Latif Jameel Poverty Action Lab*, J-PAL Advanced Course, Paris, 2012.

⁴⁰ Julia Brown, Lucia Goin, Nora Gregory, Katherine Hoffmann, and Kim Smith, *Evaluating Financial Products and Services in the US: A Toolkit for Running Randomized Controlled Trials*, New Haven, Conn.: Innovations for Poverty Action, 2015.

in addition to randomization, known as “quasi-experimental methods,” which have allowed development practitioners and researchers to apply these methods broadly.⁴¹

Not all activities require impact evaluations, and the decision to implement an impact evaluation will depend on the activity level of priority and resource intensiveness, as well as the importance of the question being answered (i.e., importance of identifying a causal relationship between activity and outcomes). Chapter Six provides more guidance on how to determine the type of evaluation to implement in practice.

Resource Intensiveness

The cost of an impact evaluation varies, but the monetary and time costs are usually considerable. Studies have estimated impact evaluation costs at as little as \$10,000 or upward of \$1 million;⁴² a cost between 5 and 7 percent of the total activity budget is not atypical.⁴³ Costs are driven by the complexity of the evaluation question that is being answered, the length of activity or evaluation implementation, the activity sample size and scope, the evaluation’s level of rigor, the purchasing power parity of the country where the evaluation takes place, and the complexity of the data collection process.⁴⁴

Impact evaluations can be time-consuming and costly because there needs to be enough time for long-term outcomes to materialize and for data to be collected at different stages. For instance, understanding whether people who participate in a training program improve their performance due to the training received would entail collecting data before, during, and after training, and then analyzing all of the data collected. This process can be time-consuming and expensive. However, the costs and length of impact evaluations can also be reduced by focusing on the most relevant questions and by reducing the need to use innovative methods.⁴⁵

These modifications would allow for considerable reductions in the time and cost of impact evaluations for the Army. For instance, consider the Army Maneuver Captains Career Course (MCCC). The main outcome of interest for an MCCC is partner nation officers performing combat tasks after course completion. A “novel” impact evaluation may involve developing a randomized experiment where some partner nation officers are assigned to participate in the MCCC while others are not, and where results on MCCC effectiveness can be generalized to all the different MCCCs. This randomized experiment may involve a large sample of officers and partner nations to produce results that can be generalized. While this experiment may be rigorous enough to be published in a scientific journal, developing a randomized experiment to assess MCCC effectiveness may not be worth the costs. Planners may choose to develop an impact evaluation that is not as rigorous as a randomized experiment, but that is faster, more affordable and institutionally viable. For example, an alternative evaluation

⁴¹ Joshua Angrist and Jorn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist’s Companion*, Princeton, N.J.: Princeton University Press, 2009; William R. Shadish, Thomas D. Cook, and Donald T. Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Boston: Cengage Learning, 2002.

⁴² Neil Buddy Shah, Paul Wang, Andrew Fraker, and Daniel Gastfriend, “Evaluations with Impact: Decision-Focused Impact Evaluation as a Practical Policymaking Tool,” 3ie International Initiative for Impact Evaluation, Working Paper #25, September 2015.

⁴³ W. K. Kellogg Foundation, *W. K. Kellogg Foundation Evaluation Handbook*, January 1, 2010.

⁴⁴ Shah et al., 2015; John Blomquist, “Impact Evaluation of Social Programs: A Policy Perspective,” *World Bank*, Social Safety Nets Primer Notes, No. 14, September 2003.

⁴⁵ Shah et al., 2015.

that could be used to estimate the impact of the MCCC is a difference-in-differences (DID) approach.⁴⁶

Misperceptions

Even though most impact evaluations can be expensive and time-consuming,⁴⁷ their costs could decrease significantly if they were prioritized and planned in the initial stages of the activities and if evaluators could rely on administrative data as much as possible. Using good-quality administrative microdata can be significantly cheaper than running the usual surveys that are used in impact evaluations.⁴⁸ In practice, most impact evaluations do not rely on administrative data and have to engage in additional data collection efforts (e.g., household surveys).⁴⁹ There is also the misperception that impact evaluations always rely on randomly assigning one group to receive the treatment and not providing treatment to another group. However, there are diverse impact evaluation methods that can be used to assess impact and that do not require randomized assignment.

How Impact Evaluation Fits with Army Current Practices

Impact evaluation provides tremendous value, but given the resources and complexity that it involves, relatively few security cooperation activities will require or justify an impact evaluation. Most impact evaluation methods tend to be data and design intensive, and the typical security cooperation activity will not support the data needs or accommodate the challenges of constructing a valid control group. Nevertheless, impact evaluation is feasible for some security cooperation activities, and it can provide answers to questions that neither performance evaluation nor monitoring is able to answer. For example, the Army could use impact evaluation to determine whether a specific Medical Civil Assistance Program (MEDCAP) improves health conditions of patients, or whether partner nation officers who participate in an MCCC are able to perform combat tasks due to the MCCC and not to something else. Impact evaluation could also estimate the magnitude of the impact—or effect—of the

⁴⁶ A DID approach estimates impact by comparing changes in partner nation officers' ability to perform combat tasks (i.e., outcome of interest) over time, between a group of officers who do not participate in the MCCC (control group) and a group of officers who participate in the MCCC (treatment group). It may be easier to implement because it does not require the use of randomization to decide which partner nation officers will participate in the MCCC and which ones will not, which may increase political viability. DID could also be implemented faster and be more affordable than a randomized impact evaluation when it relies on administrative data. Planners could collect administrative data on the outcome of interest through applying tests to partner nation officers in the treatment and control groups before and after the MCCC takes place.

To calculate impact through DID, evaluators have to calculate two differences: first, the difference in outcomes before and after the MCCC for those in the treatment group, and second, the difference in outcomes for those in the control group before and after the MCCC takes place (B-A). Finally, impact is calculated by estimating the difference between the first and second differences just mentioned. The result would be the impact of the MCCC on the outcome of interest (i.e., the ability of partner nation officers to perform combat tasks). DID may be more flexible than randomization, but it requires more assumptions to be properly applied. When these assumptions are not satisfied, evaluators cannot ensure the impact will be the result of the MCCC, because other external factors could also be causing the impact. The main assumption required for a DID approach is for treatment and control groups to be affected in the same way by factors external to the MCCC over time. For more information on this methodology, see Gertler et al., 2011.

⁴⁷ Laura Rawlings, "Who's Afraid of Administrative Data? Why Administrative Data Can Be Faster, Cheaper and Sometimes Better," *Development Impact Blog*, World Bank, June 26, 2013.

⁴⁸ Rawlings, 2013.

⁴⁹ Administrative data are often of poor quality, have insufficient coverage, or are not sufficiently relevant to the main questions about effectiveness.

Table 5.1
Comparison of AM&E Approaches

Approach	Benefits	Costs	Timeline	Operational Demands
Initial assessment	- Identifies needs and gaps - Drives activity decisions	Low to medium	Short	Varies: basic desk review to site visits
TOC	- Provides logic for how an activity accomplishes desired goals - Basis for any other M&E approaches	Low	Short	Ability to conduct staff interviews and revision of administrative documents
Monitoring	Provides real-time data to improve current activity implementation	Low to high	As long as the activity duration	- Coordination between AM&E and implementers required - Ability to collect real-time data (i.e., special software might be required)
Performance evaluation	Establishes whether activity implementation, process, and delivery are conducted as planned and according to the TOC	Medium	Short to medium	Ability to collect data and revision of administrative documents
Impact evaluation				
Randomized evaluations	Most rigorous approach. Addresses causality by isolating impact of activity from other factors	Usually high	Long	Ability to randomize required
Nonrandomized methods	If properly applied, it isolates impact of activity from other factors	Medium	Medium	- High technical skills required - Administrative data required, and potentially for data collection

SOURCE: RAND analysis.

MEDCAP on patients' health, and the MCCC on partner nation officers' capability to perform combat tasks.

Table 5.1 presents a summary of the methods we have covered, and the most typical benefits, costs, timeline, and operational demands for each approach. This table is useful for understanding the relative characteristics for each approach, but in practice, benefits, costs, timeline, and operational demands may vary depending on each specific activity.

Qualitative and Quantitative Methods for Evaluation

We now describe the methods for data collection and analysis that are used in evaluations. These methods can be broadly categorized as *qualitative* and *quantitative*, the use of nonnumeric and numeric data, respectively. Performance and impact evaluations may rely on both quantitative and qualitative methods for estimating program effectiveness, and these methods can be used on their own or as complements. Which method to use depends on the type of evaluation question the researcher wants to answer, the type of data that can be collected, and the available resources, human and monetary.

Qualitative methods rely on words or images rather than numbers,⁵⁰ and they provide contextual evidence that explains behaviors of participants and complements the data obtained

⁵⁰ Gertler et al., 2011.

through quantitative methods.⁵¹ There are multiple methods of conducting qualitative evaluations, including surveys, observation, ethnographic studies, interviews, and focus groups. Two of the most commonly used qualitative methods are focus groups and interviews, and we will briefly describe them here (for more resources on qualitative methods see Appendix G):

- Interviews are verbally administered questionnaires. The most common types are structured, semistructured, and unstructured interviews, depending on how much control the interviewer has on the questions that are being asked and on the flow of the conversation. Interviews are used to collect information about the respondents' beliefs, opinions, behavior, experiences, and descriptions of occurrences.
- Focus groups involve a small number of participants (usually six to ten) who are all asked the same open-ended questions by the researcher. Participants usually answer the questions as a group and may even start a discussion among the group.

Quantitative methods collect and analyze data in the form of numbers, and they are often combined with rigorous statistical and econometric analyses that allow the evaluator to estimate impact. Quantitative surveys tend to involve closed-end questions in which responses can be coded as categories or numbers.

While impact evaluations should involve rigorous quantitative methods, they can also draw on complementary qualitative methods. Performance evaluations, on the other hand, can be either quantitative, qualitative, or a combination of both. Evaluations that combine both qualitative and quantitative methods are known as mixed method evaluations.

Using mixed methods can increase the validity and credibility of evaluation findings⁵² and provide a better understanding of results. For instance, consider an evaluation of an Army MCCC, where the main outcome is partner nation officers performing combat tasks after course completion. A quantitative evaluation may involve performing tests on officers to assess whether they are able to perform combat tasks after completion. A qualitative evaluation may involve conducting interviews with both officers and course providers to understand their opinions about the course and its effectiveness, and to identify opportunities for course improvement. Therefore, the results from the tests would be complemented by contextual information from officers and course providers, and the mix of methods would also allow for result triangulation.

RAND's evaluation of the Commander's Emergency Response Program (CERP) in Afghanistan is an example of a recent mixed method evaluation.⁵³ Semistructured interviews (qualitative analysis) with military officers were conducted to establish CERP's effectiveness from the perspective of its implementers, while quantitative analysis through geospatial analytic methods was used to establish the relationship between the program and its outcomes, including economic and agricultural activities and population movement. For more informa-

⁵¹ Vijayendra Rao and Michael Woolcock, "Integrating Qualitative and Quantitative Approaches in Program Evaluation," in *The Impact of Economic Policies on Poverty and Income Distribution: Evaluation Techniques and Tools*, 2003.

⁵² Michael Bamberger, "Introduction to Mixed Methods in Impact Evaluation," *InterAction*, Impact Evaluation Notes, No. 3, August 2012.

⁵³ Daniel Egel, Charles P. Ries, Ben Connable, Todd C. Helmus, Eric Robinson, Isaac Baruffi, Melissa A. Bradley, Kurt Card, Kathleen Loa, Sean Mann, Fernando Sedano, Stephan B. Seabrook, and Robert Stewart, *Investing in the Fight: Assessing the Use of the Commander's Emergency Response Program in Afghanistan*, Santa Monica, Calif.: RAND Corporation, RR-1508-OSD, 2016.

tion on references to qualitative and quantitative methods, see “Further Reading Suggestions on AM&E and Impact Evaluation Practices” at the end of this report.

Importance of Selecting a Good Partner

Partnerships between AM&E planners and partner nations have always been a core feature of development programs.⁵⁴ They are defined as “a collaborative relationship between entities to work toward shared objectives through a mutually agreed division of labor.”⁵⁵ Army security cooperation usually prioritizes partnerships with other nations for strategic reasons, but AM&E requires partnerships that may surpass the level of collaboration regularly observed with strategic partners. Conducting AM&E may require active involvement in data collection or analysis by the partner nation, depending on the security cooperation activity and evaluation type. The Army should engage in *ambitious* AM&E only with strong partners, those that understand the benefits and costs associated with AM&E, and that are willing to share these benefits and costs.

We recommend considering the following characteristics, based on Glennerster,⁵⁶ when assessing partner involvement and interest in security cooperation AM&E. A good AM&E partner usually

- shares a desire to learn whether the program is attaining its expected objectives, and puts effort into attaining them
- is keen to adjust specific components of the activity—in case it is needed—that may allow proper activity monitoring and evaluation
- is willing to put time and effort into conducting proper AM&E and sharing information that allows monitoring performance or establishing effectiveness
- has sufficient expertise implementing activities from which it may be possible to observe or measure some success or positive impact
- is capable of implementing AM&E systems successfully, or is willing to learn from AM&E planners on how to implement them.

Working with partners to implement successful AM&E also depends on planner characteristics. Ideally, planners will have the

- flexibility to establish measures of performance and effectiveness that are relevant for the partner nation, and the ability to design these measures in close collaboration with the partner nation
- willingness to adapt the activity to fit the partner nation’s AM&E interest
- disposition to share expertise on AM&E methods with the partner nation’s staff in a way that might support mutual AM&E activities, or improve activity design and implementation
- capacity to deploy AM&E technical experts on the ground in order to have full-time presence in the location where the activity is being implemented, monitored, or evaluated, and ensure frequent communication with partner nation representatives.

⁵⁴ Partnerships Group—Strategy and Resource Management, “Partnerships for Development: Proposed Actions for the World Bank, a Discussion Paper,” *The World Bank*, May 20, 1998.

⁵⁵ Andres Liebenenthal, Osvaldo N. Feinstein, and Gregory K. Ingram, “Evaluation and Development: The Partnership Dimension,” *World Bank Series on Evaluation and Development*, Vol. 6, 2004.

⁵⁶ Rachel Glennerster, “The Practicalities of Running Randomized Evaluations: Partnerships, Measurement, Ethics, and Transparency,” in *Handbook of Field Experiments*, J-PAL, 2016.

Conclusion

International development organizations face AM&E needs and challenges similar to those of the Army in terms of accountability and learning from experience. To meet those needs, the development community has developed robust, sophisticated methods for monitoring and evaluating activities. The three major components of AM&E relevant to Army security cooperation are monitoring, performance evaluation, and impact evaluation. Looking across these three components, we identified general lessons and guidance for the Army:

- Multiple development organizations, including multilateral and U.S. institutions, use AM&E frameworks for activity assessment, learning, and accountability purposes. Some development agencies are also engaging in efforts to harmonize development indicators that allow comparability across agencies.
- Before implementing any monitoring and evaluation, all activities should have an activity-level TOC that is consistent with the overall organizational strategy.
- Some form of monitoring is always desirable. Monitoring enables oversight, supports transparency, and promotes awareness within and across activities. It is a continuous process that allows improvement of activity design and implementation while in action.
- Performance evaluation provides information that allows learning about how activities are being carried out and how they can be improved. Performance evaluations are common in the development community.
- Impact evaluation helps determine whether and how activities are working to achieve the desired goals and objectives. Impact evaluations are the only type of evaluation that allow establishing a causal link between the activity and the outcomes of interest. Impact evaluations are very common in development community efforts. We expect impact evaluation can be a valuable tool for evaluating Army security cooperation, but that it will not be appropriate for most security cooperation activities.
- The decision between a performance evaluation and an impact evaluation depends on activity characteristics such as activity level of priority, resource intensiveness, and importance of the question being answered (i.e., whether the Army is trying to establish a causal relationship or whether it is trying to understand how an activity is being carried out).
- Qualitative methods provide valuable information to understand respondent behavior and contextual evidence, and measure program effectiveness. They are useful on their own or as complements to quantitative methods.
- Using mixed methods—both qualitative and quantitative methods for evaluation—increases the validity and credibility of evaluation findings, and provides a better understanding of results.⁵⁷
- The AM&E framework should be incorporated into the activity life cycle and act as a feedback loop that supports continued improvement and accountability.

⁵⁷ Bamberger, 2012.

Framework for Assessing, Monitoring, and Evaluating Army Security Cooperation

The previous two chapters outlined the goals for security cooperation AM&E and identified good practices from the international development community. This chapter translates those goals and lessons into a framework that is relevant and actionable for ASCC security cooperation planners. To do so, we identified a set of implementation principles. Framework components should be

- relevant for Army security cooperation stakeholders, and in particular for ASCC planners, who are the critical link in the Army's security cooperation planning chain
- generalizable across different GCCs, different objectives, and different activities
- consistent with the emerging OSD guidance, as captured in the DoDI on Assessment, Monitoring and Evaluation Policy for the Security Cooperation Enterprise¹
- designed to enable both accountability and learning.

These components are designed as generalizable templates that are easy to modify according to context or need. These components are meant to serve as living documents that can be modified and updated over time, enabling security cooperation planners to plan for and track security cooperation LOEs. The framework can enable accountability—for planners, implementers, and stakeholders—and institutionalize a learning process within Army for the entirety of a project or program's life cycle.

ASCC Planning for Security Cooperation

ASCCs serve as the primary interface between GCCs and the department of the Army and are responsible for the administration and support of all Army forces assigned or attached to the GCC or under its operational control. The Army integrates and synchronizes its security cooperation planning by, with, and through the ASCCs both internally (within the Army) and externally (with U.S. embassy country teams and partners).² In particular, the Security Cooperation Section located within the ASCC's movement and maneuver (G-3) functional

¹ Office of the Secretary of Defense, 2016.

² Department of the Army, *Army Strategic Guidance for Security Cooperation: An Enduring Mission for "Prevent, Shape, Win,"* Washington, D.C.: Army Doctrine Reference Publication Headquarters, August 2014a.

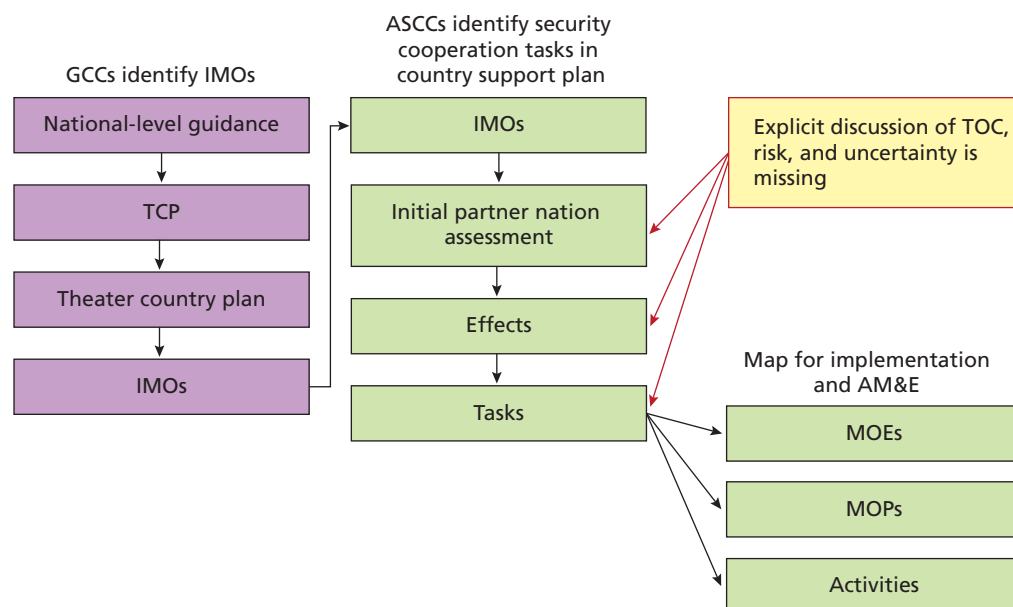
cell plans, coordinates, and provides staff supervision over all international engagement and security cooperation activities, including exercises and training.

As Figure 6.1 describes, ASCC security cooperation planning flows from national guidance and GCC strategic objectives, as laid out in GCCs' TCPs and country plans. Once GCCs identify IMOs to be achieved in theater, ASCCs plan and execute security cooperation tasks as part of country support plans to achieve GCC IMOs. Multiple security cooperation activities are often undertaken to meet each security cooperation task.

For each IMO, ASCC country support plans provide the ASCC's initial partner assessment, identify the effects that need to be created to meet the IMO, and lay out the tasks the ASCC will undertake to create the desired effects. Based on the specified tasks, ASCC planners must identify relevant activities for each task, along with corresponding MOPs and MOEs. Activity-level outputs, outcomes, MOPs, and MOEs are explained in more detail later in this chapter.

The connection between IMOs and the initial partner nation assessment is intuitively logical, if less easily managed (in the next section, we discuss strategies for matching ASCCs' country assessments with emerging OSD guidance). The weakest link in the current system is a gap between ASCCs' assessment of what partner countries need and their identification of whether activities are producing the desired effects needed to accomplish IMOs. Fundamentally, specifying desired effects does not mean that the Army will be able to achieve these effects and IMOs. In addition, the current structure does not explicitly identify and assess the risks and uncertainties involved in creating the desired effects. As a result, planners using the current structure are hampered in their abilities to identify whether activities create the desired outputs and outcomes, in general, and to develop MOEs and MOPs in particular.

Figure 6.1
ASCC Security Cooperation Planning Process



SOURCE: RAND analysis.

RAND RR2165A-6.1

Drawing on the development community's experiences, TOCs represent a tool that maps the paths to achieve desired effects and impacts (i.e., outcomes and IMO), and to identify risks and assumptions associated with each activity. Moreover, OSD has included a requirement that GCCs develop TOCs for theater-level security cooperation planning. For ASCC planners, developing TOCs to link IMOs to Army security cooperation activities will provide an opportunity to tie ASCC activities into the emerging GCC security cooperation planning process. Later in this chapter, we present a generalizable template for a security cooperation TOC.

ASCC Initial Country Assessments

Emerging OSD guidance states that initial assessments “describe host nation willingness and propensity to implement and sustain assistance, improve institutional capacity, and build capabilities in the context of country or other relevant objectives, and identify requirements, gaps, and potential risks.” If OSD's emerging guidance is adopted, GCCs will be required to undertake initial country assessments prior to all significant security cooperation initiatives at the country level. We expect that the GCCs in part will draw on ASCC expertise to complete these assessments.

More generally, initial country assessments are a key input for ASCC security cooperation planning. To accomplish GCC objectives, planners need to understand their partners' capabilities, vulnerabilities, and preferences. In turn, ASCC country assessments inform GCC strategies and priorities, creating a feedback loop from country assessment to GCC strategies, through security cooperation tasks. A partner's changing capabilities, characteristics, and strategic context will affect what security cooperation tasks are valuable and feasible. A country assessment also examines how willing and able a partner is to participate in the given activity set and how permissive the environment is for monitoring and evaluation. Country assessments provide the foundation for activity-specific initial assessments, which will be discussed in the following section on activity-specific TOCs.

All ASCCs already have procedures in place for conducting country assessments. These assessment procedures have been tailored to meet their region-specific needs within their resource constraints. We expect that ASCCs will not need to substantially alter their country assessment procedures to align with emerging OSD guidance, but recommend that ASCCs examine their current procedures for opportunities to address disconnects between the information they are currently collecting and information requests that may emerge from the GCCs.

Emerging OSD guidance indicates there are several dimensions that need to be included in a partner nation country-level initial assessment. ASCC planners need to assess partners'

- security environment
- objectives
- capabilities
- characteristics
- will
- previous engagements, particularly security cooperation activities.

Although planners should assess all of these dimensions, the depth of knowledge required for an assessment should depend on country priorities—not all partners need to be assessed

Table 6.1
Information Sources for Partner Assessments

Partner Nation Assessment Component	Information Sources for Minimal Coverage	Information Sources for Comprehensive Coverage
Security environment	Open-source reporting	Deeper intelligence reporting
Objectives	U.S. government reports on partner nation objectives (security cooperation office/defense attaché, GCC, DoS)	Direct coordination with partner nation (e.g., U.S. Army North coordination with Mexico)
Capabilities (DOTMLPF-P)	U.S. government and open-source reporting (e.g., intel and embassy reporting, Janes)	Functional area-specific assessment (e.g., RAND Security Cooperation Cooperation Continuum Assessment Framework)
Characteristics (political, military, economic, social, infrastructure, and information [PMESII])	Very minimal: U.S. government and open-source reporting (e.g., Economist Intelligence Unit) Moderate: multidimensional quantitative scorecards (e.g., political and military compatibility)	Country-specific, qualitative and quantitative SME analysis based on structured observation and interaction
Will	GCC assessment	Objective-specific ASCC assessment
Previous engagements	G-TSCMIS observations, AARs	Activity-level data repository for planning, execution and AM&E data, linked to ASCC tasks

SOURCE: RAND analysis.

with the same level of effort. For a non-high-priority country that is stable, relatively minimal assessment may be sufficient. In contrast, high-priority countries, or countries that are undergoing periods of flux or uncertainty, will require more in-depth assessment. Information sources for lower-priority countries can be primarily open-source documents that are accessible to most ASCC security cooperation planners, while more intensive assessments may require information that is closer held or privileged in some way. For these partners, ASCC planners may want to work with DoS country teams.

The nuance and depth of understanding required for each country assessment will depend on several factors, including the partner's importance and the anticipated volatility of the operational environment. For regional partners that the United States considers critical or key, a more detailed country assessment is appropriate. Partners that are undergoing major transitional periods, whether that transition is political, social, or economic, will also require more frequent or more detailed partner assessments. Table 6.1 provides an overview of potential information sources for each assessment dimension, ranging from minimal to comprehensive coverage. Appendix C discusses each of these dimensions and potential information sources in more detail.

TOC for Army Security Cooperation Activities

In Chapter Four, we reviewed the emerging OSD guidance for security cooperation and noted that the IDD that GCCs will need to develop must include TOCs. In Chapter Five, we introduced TOCs as a relevant tool for security cooperation AM&E, and we mentioned that activity-level TOCs can be used to map how security cooperation activities are expected

to achieve specified outcomes. TOCs are critical for the Army because they ensure planners define and map the causal link between activities and their effects, which in turn supports activity implementation and AM&E.

Applying TOCs to ASCC activities does not require substantial changes to ASCCs' current process, but it will require a mindset change that explicitly recognizes the uncertainty that exists in the causal assumptions linking activities to desired outcomes. This section describes how TOCs can be developed for Army-specific activities, the TOC components, and the steps involved in putting them together.

Developing TOCs for Army Security Cooperation

The first step in developing a TOC is to identify the goal to be achieved; in this case the goal is the IMO. GCCs generally assign IMOs to ASCCs through their TCPs and country plans (see Figure 6.1). IMOs also represent the objective at the end of an ASCC's TOC. IMOs are long-term outcomes or impacts that are expected to result from security cooperation activities, but they are still "intermediate" because they represent steps toward achieving the GCC's strategic objectives. The next step in developing a TOC, after identifying the IMO, is to identify a key security cooperation activity that may help achieve that IMO. The TOC will then help map the logic from activity implementation to IMO achievement.

To illustrate how to develop TOCs, we present three TOC exemplars in Appendix D. We selected three example activities that represent different levels of uncertainty embedded into program design and have different data collection and analysis limitations:³

1. Medical Civil Assistance Program (MEDCAP)
2. Maneuver Captains Career Course (MCCC) at Fort Benning, Western Hemisphere Institute for Security Cooperation (WHINSEC)
3. AFRICOM-led multilateral exercise (Southern Accord 2016).

These example activities capture a range of features that make developing a TOC relatively more or less challenging. The features include whether the activity occurs only once or is a recurring activity (see Annex E2); whether the activity has a track record of successful implementation versus whether it is being implemented for the first time or in a new context; and whether it is easy and feasible to collect data before, during, and after activity implementation (i.e., whether MOPs and MOEs are straightforward to identify). Table 6.2 provides the overall rating of each activity across the three dimensions.

We rate MEDCAP as "easy" because it is a recurring activity that is likely to be carried out again in the future with a relatively similar structure and design. U.S. Army Africa has experience implementing MEDCAPs successfully, and previous lessons have been incorporated into planning. Moreover, MOPs and MOEs are straightforward to identify and data can be collected relatively easily (i.e., medical care is a common, widespread activity). We also use the MEDCAP example throughout this section to illustrate the steps needed in a TOC.

³ These examples represent real security cooperation activities that the Army undertakes and are based on our collection of press releases from Army events, course descriptions, and news articles. However, we emphasize that they are *notional examples*. Where we lacked specific information to discuss inputs, for example, we inserted reasonable approximations. See, for example: Tiffany DeNault, "Army Civil Affairs Provides Medical Care to Djibouti Village," May 21, 2016; Western Hemisphere Institute for Security Cooperation, *Course Catalog 2017–2018*, March 17, 2015, p. 11; "Southern Accord 16," *U.S. Army Africa*, August 1, 2016.

Table 6.2
Ease of Producing a TOC

Dimension	MEDCAP	MCCC	Southern Accord
Is it recurring?	✓	✓	✗
Does it have a track record of successful implementation?	✓	✓	✗
Are MOPs and MOEs straightforward to identify and can data be collected relatively easily?	✓	✗	✗
Difficulty for establishing a theory of change	Easy	Medium	Hard

SOURCE: RAND analysis.

The MCCC is rated as “medium” because it is a recurring activity that will likely be carried out in the future with a similar structure, but success requires strong assumptions and tracking progress is more difficult. WHINSEC keeps track of student and instructor performance for all sessions, making it possible to show a track record of successful implementation, but the course system is not set up to track students after course completion. This hinders evaluation of MOEs and attribution of specific outcomes to MCCC implementation.

The notional Southern Accord 2016 activity implemented in Malawi is the hardest activity for which to develop a TOC. We categorize it as *nonrecurring* because although the exercise occurs annually its location changes every year and the design is adapted for the host nation; the fundamentals of the activity change sufficiently to render each year’s activity distinct. The changing context and design of each Southern Accord also makes it difficult to prove a track record of successful implementation and performance. In addition, it is difficult to collect data after activity completion and establish a causal relationship between the activity and the desired outcomes.

In the following section we briefly describe the steps required to establish a TOC; our examples focus on MEDCAP. However, readers may look at Appendix D for more detail on what a full TOC looks like for MEDCAP, WHINSEC, and Southern Accord.

Steps for Establishing an Activity-Level TOC in Practice

Planners implementing an activity-level TOC should perform the following steps:

1. Identify IMOs
2. Decide on activities that are key to achieving IMOs
3. Develop a causal logic for the TOC, which includes how IMOs connect back to activities and inputs
4. Conduct an initial assessment for activity planning
5. Establish measures for outputs (MOPs) and outcomes (MOEs)
6. Describe underlying assumptions.

Identify IMOs

If the IMO is not specified by the GCC, ASCC planners should identify the IMO by answering “What are the intermediate goals that have to be achieved to accomplish GCC strategic objectives?”

For example, consider the following GCC strategic objective: *Develop medical interoperability with coalition partners*. Planners could identify the following IMO as an intermediate

objective to the GCC strategic objective: *Improving partner nation's preventive health care system*. Additional examples for IMOs are shown in Appendix D.

Decide on Activities to Achieve IMOs

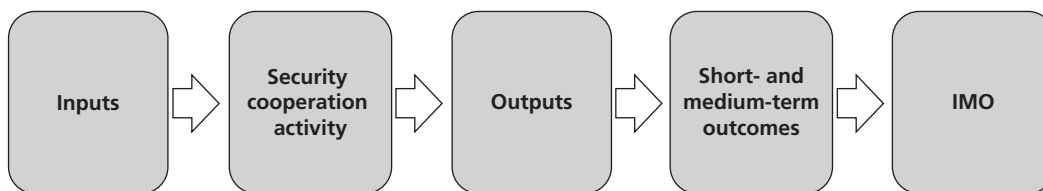
After identifying the IMO, planners need to decide on the activity or group of activities that have to take place to achieve the IMO. In a TOC, activities are actions implemented to achieve IMOs. There can be multiple activities in action to achieve one single IMO or one activity supporting multiple IMOs. The IMO we had identified was *improving partner nation's preventive health care system*. An activity that may lead to this IMO could be MEDCAP, specifically: *U.S. physicians provide medical training to partner nation physicians*. For our purposes, we will use only one activity to describe MEDCAP.

Develop TOC Causal Logic

After the planner has thought about whether she might be able to achieve the intended IMOs, she must think about the *pathway* or logic that leads to the intended IMOs.⁴ The pathway connects IMOs with outcomes, outcomes with outputs, outputs with activities, and activities with inputs—terms we further define in the upcoming paragraphs. TOCs develop a causal logic from inputs to IMOs, as illustrated in Figure 6.2. It is important to note that although the logic chain in Figure 6.2 starts with inputs and ends with IMOs, planners work through the TOC from the IMOs they want to achieve and work backward to identify the links in the chain that must be forged to achieve an IMO.

The first step in establishing the TOC causal logic is to identify the short- and medium-term **outcomes** that might lead to IMOs. Outcomes are achieved after the participants in an activity use or apply outputs;⁵ hence, they are the short- and medium-term effects from activity implementation. For instance, in our MEDCAP example, the IMO was *improving partner nation's preventive health care system*. Its corresponding short- and medium-term outcome would be *trained partner nation physicians are able to provide high-quality patient consultations and diagnostics on their own after MEDCAP*. The MEDCAP outcome happens only after trainees apply the MEDCAP training (i.e., after trainees apply the MEDCAP outputs). The measures that are used to assess whether outcomes are being produced as expected are called MOEs.

Figure 6.2
Causal Logic of an Activity-Level TOC



SOURCE: RAND analysis.

RAND RR2165A-6.2

⁴ Center for Theory of Change, n.d.

⁵ Gertler et al., 2011.

The second step is to connect **outputs** to outcomes. Outputs are tangible, short-term results or services that activities accomplish or deliver.⁶ Activity participants use outputs to achieve outcomes. For instance, in our MEDCAP example, an output would be the *completed provision of the two-day medical training to partner nation physicians*. If partner nation physicians use and benefit from the two-day training, then the outcome would be *trained partner nation physicians are able to provide high-quality patient consultations and diagnostics*. The indicators that are used to assess whether outputs are being produced as expected are called MOPs.

A key difference between outputs and outcomes is that security cooperation planners have control over outputs, which can be tracked or reported on relatively easily.⁷ However, planners have much less control over outcomes because external factors can affect whether activity beneficiaries use or apply the outputs. For instance, partner nation physicians may complete the full MEDCAP training provided by U.S. physicians (i.e., the output is the completed training), but this output may not be sufficient for partner nation physicians to provide high-quality consultations and diagnostics on their own (i.e., the outcome). Being able to provide high-quality consultations independently is also affected by external factors outside the control of the security cooperation planner, such as partner nation physicians having adequate medical equipment or partner nation physicians actually learning the content of the training. As a result, it is much easier for security cooperation planners to track whether MOPs are being produced as expected than it is to assess MOEs. This is a key reason why, historically, AM&E has been concentrated in monitoring only (which focuses on MOPs), and has disregarded evaluation methods (which estimate changes in MOEs). As mentioned in Chapter Five, the use of evaluation methods has grown exponentially in recent decades.

The third step in our TOC is connecting **activities** to outputs. Activities are actions that take place using resources (inputs) to produce outputs. In our MEDCAP example, our key activity is *U.S. physicians provide medical training to partner nation physicians*. This activity supports outputs because U.S. physicians need to provide medical training to produce the outputs.

The fourth step is to connect **inputs** and activities. Inputs are all the resources required for activities to take place.⁸ Inputs are most commonly monetary resources, human resources, or infrastructure needs. In our MEDCAP example, the main activity was the provision of medical training from the U.S. physicians to partner nation physicians, and the inputs required include U.S. and partner nation physicians, medications, medical equipment, and clinical facilities. In this example, inputs support activities because U.S. physicians are able to provide medical training through the use of medical equipment, medications, and facilities.

Security cooperation achieves its goals by transforming inputs into IMO through this causal, logical chain. Once security cooperation planners identify inputs, activities, outputs, outcomes, and IMOs, planners can connect them in a TOC that describes the causal relationship between them. An example of this causal logic is illustrated in Figures 6.2 and 6.3. As planners develop TOCs for recursive activities, subsequent planning processes can begin with prior activities' TOCs, updating the assumptions embedded in the TOC based on lessons learned from previous activities.

⁶ Gertler et al., 2011.

⁷ An exception would be when recipients of an activity fail to participate in the activity, often called "noncompliance." In the case of noncompliance, the activity may not produce outputs as planned.

⁸ Gertler et al., 2011.

Conduct Initial Assessment for Activity Planning

An initial assessment for activity planning requires examining the partner nation and the activity before the activity gets implemented to decide whether the activity is needed and feasible. We outline specific steps to conduct an initial assessment for activity planning. These may constitute a more intensive assessment than the ongoing country assessment described in the previous section, particularly for a nonpriority partner, but they focus on similar issues designing more effective activities and mitigating implementation risks. For activity planning, an initial assessment should⁹

- analyze capabilities and characteristics of the partner nation
- examine potential risks and indirect effects derived from activity implementation¹⁰
- describe information that is relevant for activity design such as type of data available, baseline and target measures, and activity objectives
- investigate contextual factors that affect activity implementation and impact (e.g., cultural, economic, political, environmental)
- review external ongoing activities or programs that could affect activity implementation
- provide any other information that could inform the initial assessment.

Appendix D provides examples on how to develop the activity-level initial assessments for the MEDCAP, WHINSEC and Southern Accord examples mentioned earlier.

Establish Measures for Outputs and Outcomes

Security cooperation planners can track and assess performance and impact by collecting information through monitoring documents (including AARs), activity initial assessments, staff interviews, and administrative data.

MOPs are used to monitor and evaluate progress toward expected outputs. As described in Chapter Five, these measures should also be SMART,¹¹ and they can be either quantitative or qualitative. Quantitative MOPs include percentages, rates, ratios, and numbers.¹² For instance, the output in our MEDCAP example was *full provision of the two-day medical training to partner nation physicians*. Examples of three quantitative MOPs would be *number of partner nation physicians registered for training*, *number of partner nation physicians that completed the training*, and *number of patients seen by partner nation and U.S. physicians*. Qualitative MOPs can measure compliance with, quality of, extent of, and level of a certain result. Two examples of qualitative MOPs would be *level of partner nation and U.S. physician collaboration*

⁹ Department of Defense, “Assessment, Monitoring, and Evaluation Policy for the Security Cooperation Enterprise,” n.d.

¹⁰ In particular, planners should try to answer the following questions: What other unexpected outcomes could result from activities? What other external factors could influence the outcome? For instance, a potential risk associated with MEDCAP provision is having U.S. physicians getting infected with a new or severe disease existent in the partner nation. An example of an indirect effect is the positive spillover effect that having healthier groups of people (i.e., patients who receive consultations directly) can have on the partner nation population as a whole. Since patients who participated in MEDCAP are now healthier and can no longer infect others with the disease, the partner nation population becomes less sick in general. Another positive secondary effect from MEDCAP is that having more trained partner nation physicians could have positive spillovers for partner nation physicians in general because those who received training may share their knowledge with others.

¹¹ McNerney et al., 2016.

¹² Saraswathi Menon, Judith Karl, and Kanni Wignaraja, *Handbook on Planning, Monitoring and Evaluating for Development Results*, New York: UNDP Evaluation Office, 2009.

during MEDCAP training, and extent to which partner nation physicians adhere to the U.S. physician recommendations on how to provide consultations. Appendix D presents examples of MOPs for MEDCAP, WHINSEC, and Southern Accord. The appendix also displays how MOPs fit in with the TOC.

MOEs monitor and evaluate progress toward expected outcomes. Similar to MOPs, MOEs have to be SMART and can be qualitative or quantitative. For instance, in MEDCAP the expected outcome was *trained partner nation physicians are able to provide high-quality patient consultations and diagnostics*. Examples of quantitative MOEs that planners could use to track progress toward this outcome are *number of partner nation medical personnel who are able to provide consultations and diagnostics independently*, and *number of patients who receive consultations and diagnostics by trained medical personnel after MEDCAP*. Examples of two qualitative MOEs include *extent to which partner nation physicians are able to provide high-quality consultations to patients*, and *overall patient satisfaction with partner nation physician performance during consultations*.

Describe Underlying Assumptions

TOCs rely on assumptions about why and how change occurs as a result of the activity.¹³ In particular, security cooperation planners need to make assumptions about how the security cooperation activity solves an underlying problem or satisfies a need, whether the context in which the activity takes place is adequate for implementing the activity, whether the timeline of the activity is long enough to observe how inputs lead to IMO, whether the individuals receiving an activity actually use and benefit from the activity, and the availability of inputs needed for an activity to take place. In our MEDCAP example, security cooperation planners could formulate assumptions around

- the demand for medical services from patients who need consultations
- the need for medical training from partner nation physicians
- the effectiveness of the training as delivered to partner nation physicians.

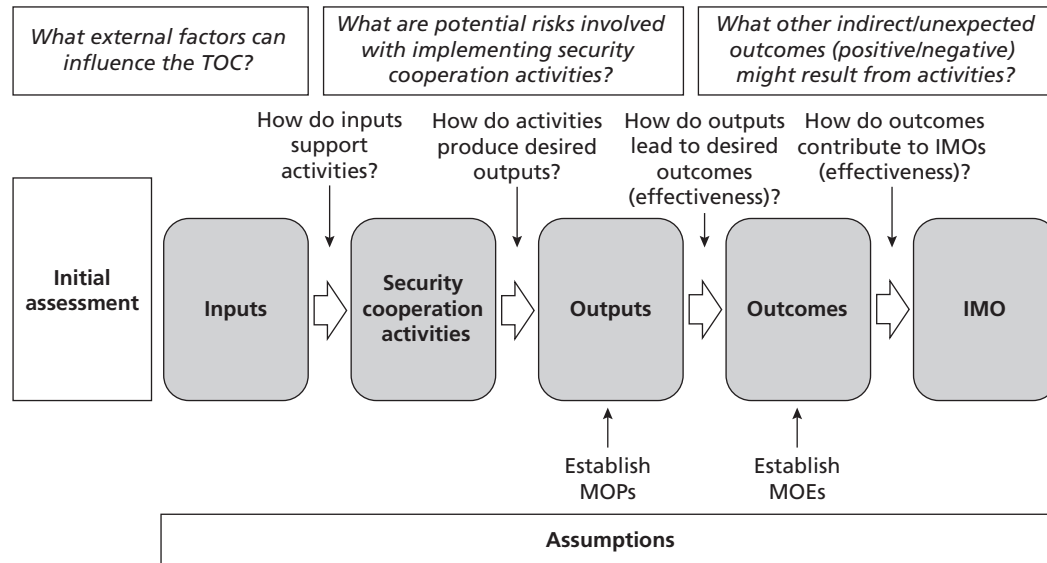
After this process, planners can build a TOC like the one shown in Figure 6.3. Appendix D presents examples of what the complete TOCs, including underlying assumptions, look like in practice and for the MEDCAP, WHINSEC, and Southern Accord examples mentioned earlier.

Changes to Current Army Security Cooperation Approach

The previous steps for establishing a TOC in practice conform well with current security cooperation planning guidance, as embodied in the security cooperation planners' course discussed in Chapter Four, and represent the mindset shift from the current Army system to a TOC approach. Currently, Army security cooperation lacks a standardized system where TOCs are established at activity inception, and where these TOCs help determine the measures of performance and effectiveness used when estimating activity contributions to campaign objectives. Army security cooperation also implements activities without systematically and consistently analyzing the underlying assumptions for each activity that gets implemented, including the analysis of the problem, the context, the timeline, the activity participants, and the availability

¹³ Isabel Vogel, *Review of the Use of 'Theory of Change' in International Development*, Review Report, United Kingdom: Department for International Development (DFID), 2012.

Figure 6.3
Components Involved in a TOC



SOURCE: RAND analysis.

RAND RR2165A-6.3

of inputs. Finally, the measures of performance and effectiveness that are used differ for activities with similar TOCs and are not comparable across the variety of DoD institutions.

Utilizing the proposed TOC framework would require systematizing and standardizing the use of TOCs, its underlying assumptions, and accompanying measures of performance and effectiveness. TOCs would need to be implemented at activity inception for all activities; activity-level initial assessments would consistently follow the same steps across all activities; and measures of performance and effectiveness would need to be harmonized across activities and DoD institutions in the longer term. These changes would allow formulating TOCs that can be used for monitoring and evaluation purposes, comparing the costs and benefits associated with activities, and identifying activities that are better contributing to DoD-wide campaign objectives.

Army Security Cooperation M&E

In Chapter Five we examined three M&E strategies: monitoring, performance evaluation, and impact evaluation. In this section we examine how these M&E strategies can be applied to the security cooperation activities in practice. All security cooperation activities require some level of AM&E, as there is value in collecting information about “how much” of an activity was done, where it was done, and who conducted the activity.¹⁴ Even basic monitoring can help influence decisionmaking or improve activity design and implementation over time. However, not all activities require rigorous M&E methods such as impact evaluation or even performance

¹⁴ Burt et al., 1997.

evaluation. The guiding principle in choosing an appropriate level of M&E is based on a cost-benefit calculation: a planner should balance the benefits of learning and documentation (to support accountability) with the costs of implementing M&E. In this section we provide basic guidance on how to achieve that balance and select the right M&E approach for each security cooperation activity.

Monitoring for Security Cooperation Activities

Monitoring should have a central role in security cooperation planning and implementation. Monitoring is useful because, by continuously tracking inputs, activities, outputs, and sometimes outcomes, it directly supports auditing, reporting, and activity course correction.

AARs can be used as part of the monitoring process to provide immediate feedback after an event takes place and allow programmatic adjustments when the activity is carried out in the future. However, previous RAND research has identified the AAR process as a gap in current security planning.¹⁵ For AARs to be useful, they should be based on the activity TOC and capture how security cooperation activities are carried out, how inputs are used, what outputs are expected to be achieved, and the level of progress toward achieving specific outputs. If after-action reporting is implemented consistently across activities, and if the reports are stored and can be retrieved, it will provide information that will be beneficial for M&E by informing midcourse correction and performance and impact evaluations.

Monitoring that is more robust and intensive than AARs is recommended when planners require stronger performance measurement and accountability mechanisms, in which case AAR tends to be insufficient. This more intensive monitoring may require more extensive data collection through surveys, focus groups, administrative data, and direct observation in addition to relying on AARs. Monitoring should be scaled up based on costs and benefits: the benefits created through the data that will be collected and used against the costs involved in data collection.

A robust monitoring system tracks measures for each activity, follows standard measures across all activities, and allows for comparability of performance and progress across activities. In this system, all activities require some form of monitoring, and the information generated is used for auditing and reporting purposes, and to promote a culture of learning and accountability. Monitoring underpins evaluation, and through improved monitoring systems, the Army could carry out high-quality evaluations and promote learning.

To understand the purpose of monitoring, consider the MCCC example, which requires more advanced monitoring than simple AARs (Appendix E and Annexes E1–E11 explain how to determine the type of monitoring needed). Monitoring would be following and recording measures quite frequently. It would always track MOPs, and sometimes track MOEs. Some examples of these measures would include the number of students trained during the course, the percentage of training content covered up to a specific date, the level of student satisfaction with the training received, the extent to which trainers teach the MCCC standardized content, and the extent to which students adhere to training recommendations during practices. Therefore, monitoring could be used by Army security cooperation planners to identify any issues with training implementation and can also be used for course correction as needed. It would be much better for planners to identify any issues with activity implementation throughout the

¹⁵ Watts et al., 2017.

implementation rather than after the activity has finished. Monitoring can also be useful to identify whether assumptions hold throughout implementation—for instance, if an assumption was made about the timeline, and the monitoring process shows that the MCCC timeline is not enough to cover all the course material. Evaluators could also use different methods to collect monitoring data, including AARs, administrative records, surveys, and focus groups, and use these data as inputs for a performance evaluation. The results from the monitoring process would be used to improve activity performance, allow midterm course correction, and inform future activity design. Finally, if different MCCCs are being monitored, then planners could use the data produced from different MCCCs to compare performance across activities.

Planners interested in learning how to classify activities into the most basic type of monitoring (AARs) or into the more intensive type of monitoring should look at Appendix E and Annexes E1–E11 for guidance. This appendix and annexes also explain the type of monitoring that is required for our three activity examples: MEDCAP, WHINSEC, and Southern Accord.

Performance Evaluation for Security Cooperation Activities

Performance evaluation can help security cooperation planners understand whether an activity is being implemented effectively and as planned. It is most useful for identifying how inputs were used, what outputs (and sometimes outcomes) were produced, and whether there were challenges in implementing the activity. Accordingly, performance evaluation is well suited to support accountability and to help improve how repeated activities are carried out over time.

The value of conducting performance evaluations for the Army will vary for each activity, and not all security cooperation activities require performance evaluations. Many security cooperation activities that are inexpensive, one-off activities will not benefit from performance evaluations. As implementation costs increase and activities are repeated over time, the benefits of performance evaluations will rise. For activity implemented over long time frames, the Army may consider conducting a performance evaluation in the middle of the activity for midcourse correction. Midcourse evaluations usually examine inputs, activities, and outputs. If an activity has been completed or enough time has elapsed since the activity was initially carried out, a performance evaluation can also examine outcomes achieved. However, this method is unable to establish whether changes in outcomes are caused by the security cooperation activity or something else (see Chapter Five for an explanation on causality). For shorter activities that are repeated frequently, a performance evaluation conducted after the activity has been completed can improve activity quality and future success.

Ideally, performance evaluations should be planned at activity inception, so that the evaluation planning, data collection, and analysis are implemented in tandem with the activity. This allows the activity and evaluation plans to be coordinated, and it allows for proper feedback and accountability mechanisms. When the Army decides to conduct a performance evaluation on an activity, it should also coordinate AARs and other monitoring efforts, which support systematic data collection that feeds into the evaluation.

Performance evaluations can be either quantitative or qualitative, but the Army would benefit more from using a mix of both methods. A performance evaluation that uses mixed methods would allow the Army to understand whether an activity is accomplishing expected outputs—usually determined through quantitative metrics—and what is the level of participant accomplishment and satisfaction with the outputs produced.

To understand the purpose of performance evaluation, consider the MCCC. A performance evaluation would analyze activity information that includes the number of officers

starting and finishing the course and examine the course content, including surveys to MCCC providers and participants to understand whether the course syllabus is aligned with expectations and the intended IMOs. A performance evaluation would also examine whether course expectations are being met by implementers and beneficiaries once the course is completed. The evaluation should review administrative data collected throughout the activity to estimate MCCC efficiency and determine areas of improvement.

Planners can rely on Appendix E and Annexes E1–E11 to determine whether a specific activity requires a performance evaluation.

Impact Evaluation for Security Cooperation Activities

Security cooperation activities seek to achieve short- and medium-term outcomes in support of IMOs. Hence the Army will be most effective in its security cooperation mission if it can say with confidence that the resources it spends on a specific security cooperation activity directly lead to the desired outcomes and IMOs. However, not all AM&E tools can provide the evidence needed to establish cause and effect. Impact evaluation provides value because it can help security cooperation planners establish whether and how specific outcomes are caused by a security cooperation activity and not by something else. This level of rigor and evidence can be costly, and not all security cooperation activities will require impact evaluation. In fact, given the cost and complexity of impact evaluation, relatively few security cooperation activities will warrant it.

The most effective way to know whether a security cooperation activity is working is to understand what would have happened had the activity not been carried out. Impact evaluation methods take this approach by comparing a “treatment” group and a “comparison” group to measure the causal effect of an activity. The treatment group is typically the set of individuals that received or participated in the activity.¹⁶ For example, for a MEDCAP activity the treated group would be those partner nation physicians who receive training from U.S. physicians. The comparison group in this example would be a set of partner nation physicians who did not participate in the MEDCAP training but *who are similar to the treatment group*. If the two groups are similar enough, then the only original difference between both groups—in statistical terms—is MEDCAP participation, and the evaluation can effectively measure the effect of MEDCAP on outcomes of interest by comparing the treatment and comparison group outcomes after MEDCAP training is complete.

Overall, most impact evaluation methods tend to be data, resource, and design intensive, and the typical security cooperation activity will not support the data needs or accommodate the challenges of constructing a valid comparison group. Therefore, for many security cooperation activities, it will not make sense for the Army to spend the resources required to carry out impact evaluation. Nevertheless, impact evaluation is feasible for some security cooperation activities. For security cooperation activities that are amenable to impact evaluation *and* where the benefit of high-quality evaluation and learning is high, the Army should strongly consider using impact evaluation. While comparing the benefits versus the costs of implementing an impact evaluation, the Army could also explore different alternatives to reduce the cost of data collection associated with impact evaluation (see Chapter Five). These alternatives include rely-

¹⁶ A treatment group could be defined by organizations or even geographic areas, rather than individuals. For example, a joint counterterrorism operation with a foreign military might be applied to specific parts of a country.

ing on administrative data as much as possible instead of engaging in more demanding data collection methods.

Similar to performance evaluations, impact evaluations are most useful when they rely on both quantitative and qualitative methods. Quantitative methods would allow establishing whether the activity had the desired effect or impact, and the magnitude of the impact. Qualitative methods would provide information about how and why this impact took place, and how the activity could be improved.

A useful example of how impact evaluation can be applied to security cooperation is the MCCC. The main outcome for MCCC is partner nation officers performing combat tasks after course completion. Outcomes cannot be captured in the number of officers who complete the MCCC, or by asking MCCC participants whether they believe the course has helped them perform combat tasks (i.e., by looking at outputs). Applied to this example, impact evaluation would allow security cooperation planners to determine whether partner nation officers are able to perform combat tasks due to the MCCC—and not to something else—and determine the magnitude of the increase in partner nation officers' capability as a result of the training.

The MCCC example also helps illustrate why impact evaluation provides distinct benefits compared with other types of AM&E. Consider a situation in which partner nation officers eligible for MCCC are already high performers in combat tasks prior to enrolling in the MCCC. Imagine also that those who decide not to enroll in the MCCC ignore course material about how to perform combat tasks and have no interest in learning on their own. If evaluators were to compare the outcomes of participants with those of nonparticipants once the MCCC is completed, they could conclude that the difference in performance between participants and nonparticipants is due only to the MCCC, overstating the MCCC's impact. An impact evaluation would compare outcomes of participants with those of a carefully selected comparison group. In this case, evaluators would compare outcomes of partner nation officers with similar characteristics (e.g., both groups perform combat tasks before the MCCC), and their only difference would be MCCC participation. As a result, the difference in the outcomes of these two groups would accurately estimate the impact of the MCCC.

Another benefit of impact evaluation is that it can help illustrate how different types of participants respond to an activity. For example, in the case of MCCC, impact evaluation could be used to measure the relative performance of participants that come in as "high performers" versus participants that come in as relatively lower performers. Impact evaluation methods can often distinguish between these "subgroups" when it comes to the causal effect of the activity.

Planners can use Appendix E and Annexes A–K to determine which activities would require an impact evaluation, given specific activity characteristics such as recurrence, resource intensiveness, and the level of existing evidence to support the causal effect of the activity on the outcomes of interest.

Decision Tree to Guide M&E Decisions

M&E strategies range from minimally resource intensive (conducting AARs) to potentially highly resource intensive (conducting impact evaluations). We recommend that all security cooperation activities involve some type of M&E, but the type of M&E used should be determined on a case-by-case basis. Which M&E strategy to adopt should reflect planners' need for learning and constraints on effective learning. Activities that occur frequently, are resource intensive, and for which little is known about the causal relationship between the activity and the desired outcomes require more rigorous or demanding M&E methods. In contrast, security

cooperation activities that occur infrequently or are one-offs, those that use few resources, and those for which there is existing evidence of effectiveness need only limited M&E. We consider when planners should adopt four different types of M&E strategies:

- AARs (only)
- Monitoring
- Performance evaluation
- Impact evaluation.

At the heart of our framework is a decision tree to guide planners' choices on the M&E method to use when considering the appropriate type of M&E for each security cooperation activity (Figure 6.4). Planners can use this to determine which is the most relevant and appropriate form of M&E: monitoring, performance evaluation, or impact evaluation. The decision tree also reminds planners of the fundamentals that need to be in place before implementing any activity-level M&E method.

As an initial check, the decision tree asks planners to confirm that the planned activity fits within ASCC security cooperation objectives and has a specified TOC (boxes 1–4). The decision tree is designed to help planners work through a series of questions to identify appropriate activity-level M&E strategies.

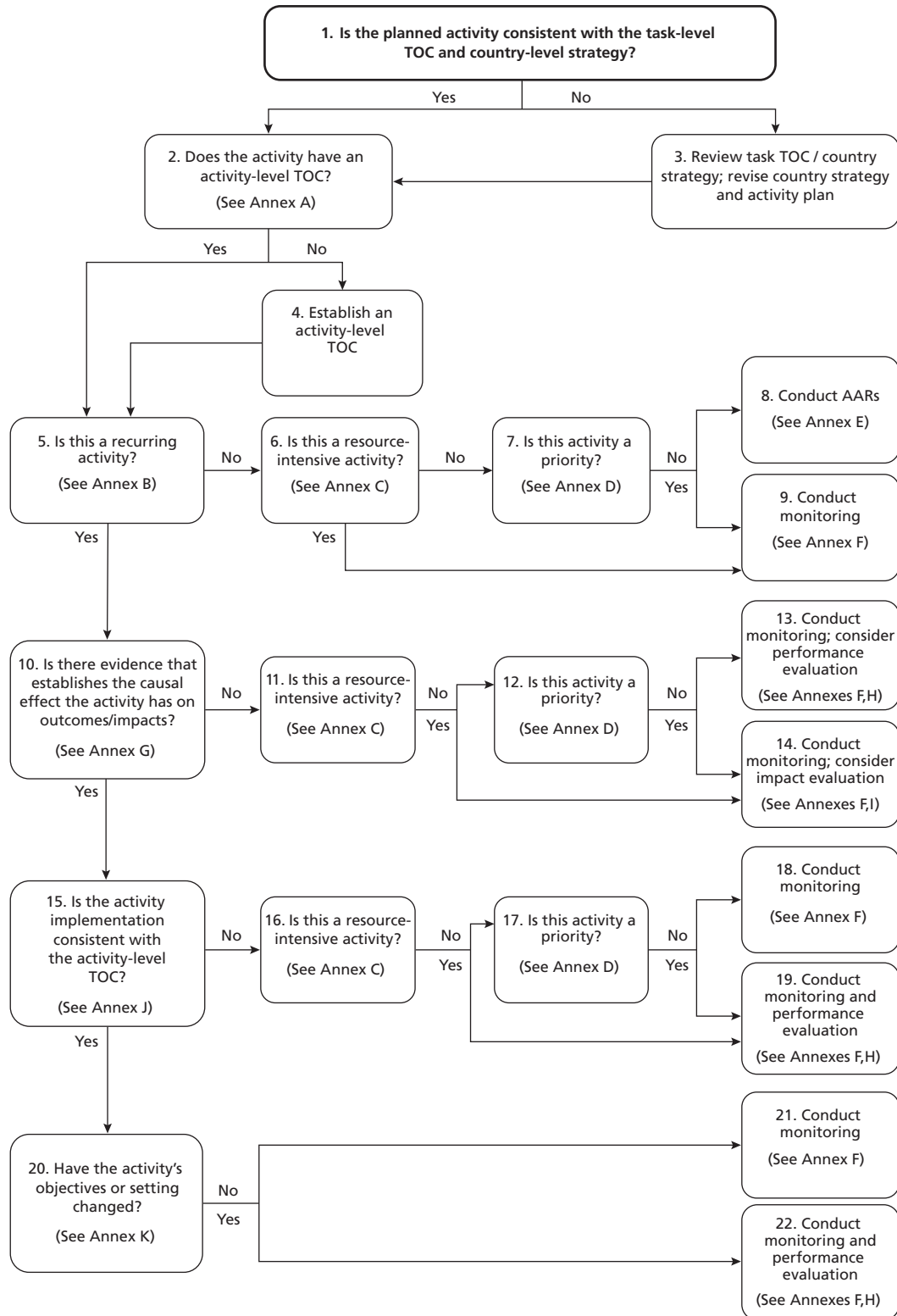
- Is the activity recurring (box 5)?¹⁷ Recurring activities provide both a greater need and opportunity for learning through M&E. An example of a recurring activity is an annual training course.
- Is there evidence that establishes the causal effect of the activity on specific outcomes (box 10)? This question highlights how uncertain planners might be about the planned activity's effectiveness. For activities with well-understood causal relationships, the need for additional learning about effectiveness might be low. Alternatively, activities that are well understood but are implemented in a different manner in which previous lessons were learned (box 15), in a different context (box 20), or to accomplish a different goal (box 20) may need additional evaluation.
- Is the activity resource intensive (boxes 6, 11, and 16)? Evaluation is often expensive. All else being equal, there is a greater need for learning about what makes activities effective when the activities are resource intensive than when they are not.
- Is the activity a priority (boxes 7, 12, and 17)? Similarly to resource-intensive activities, all else being equal, there is a greater need for learning about what makes activities effective when the activities are ASCC priorities than when they are not.

For each question, we have produced an annex to the decision tree that explains what the question is asking and defines core concepts. These are included in Appendix E. Taken together, the decision tree and Appendix E should provide enough information for planners to navigate the M&E selection process, but planners will benefit from additional information (much of which is found in the body of this report) and potentially training.

Figure 6.5 presents a notional M&E strategy for Southern Accord, a multilateral exercise held by U.S. Army Africa. We discuss this example in more detail in Appendixes D and E. For

¹⁷ By *recurring*, we mean that the activity is repeated over time while its structure and design remain stable.

Figure 6.4
Activity-Level M&E Decision Tree

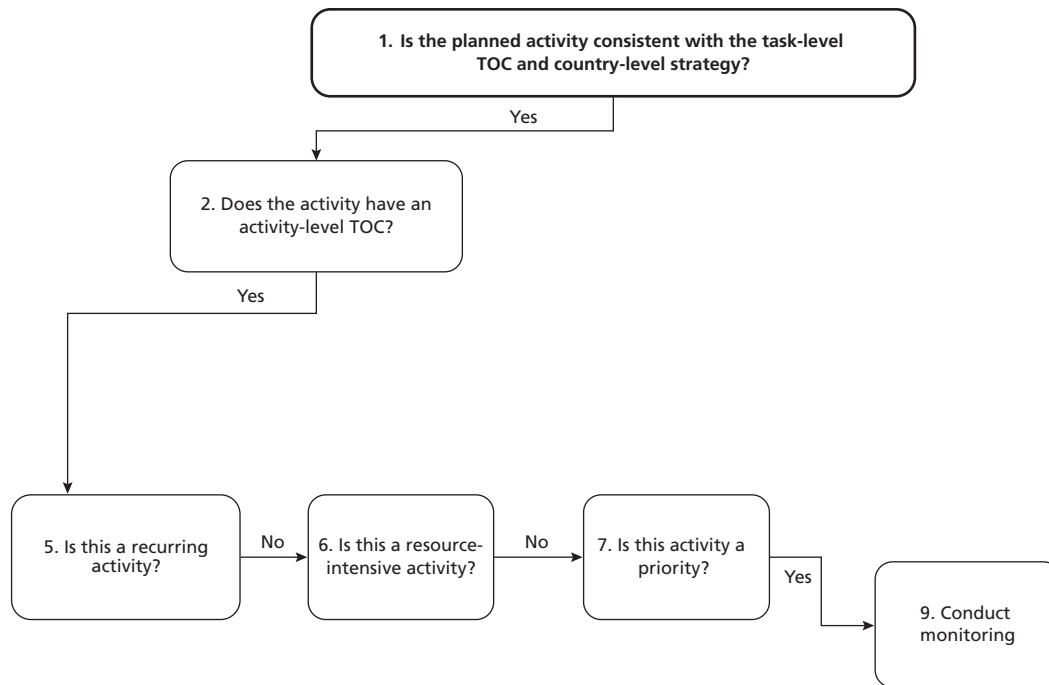


SOURCE: RAND analysis.

NOTE: Annexes referenced in this figure are included in Appendix E.

RAND RR2165A-6.4

Figure 6.5
Exemplar Activity-Level M&E Decision Tree



SOURCE: RAND analysis.
 RAND RR2165A-6.5

this activity, we assumed that planners have developed an activity-level TOC and that Southern Accord matches with USARAF's strategic objectives (boxes 1 and 2). Three decision nodes guided the notional path we outline for Southern Accord. First, is this a recurring activity (box 5)? Southern Accord is a recurring series of multilateral activities, but each one is relatively unique in its execution. As a result, we expect that most planners will view Southern Accord as a nonrecurring activity. In contrast, most training courses would be recurring activities. Second, is this activity resource intensive (box 6)? Historically, Southern Accord is a command post and tabletop exercise, which are generally considered medium-cost training exercises—they are much more resource intensive than subject matter exchanges, but are dwarfed by an undertaking such as Pacific Engagement. Third, is this activity a priority? From discussions with U.S. Army Africa planners for previous RAND research, Southern Accord was identified as a priority activity.¹⁸ Southern Accord, and its associated African Land Forces Summit, provides a valuable forum for U.S. Africa and regional leaders to engage. Following the guidance of these decision nodes, we recommend that an activity such as Southern Accord incorporate a monitoring strategy in the activity implementation. (Appendix E presents additional activity pathways through the decision tree.)

We used this framework to identify M&E strategies for four notional security cooperation activities held with Bandarian Land Forces. As can be seen in Table 6.3, the types of M&E strategies recommended reflect differences in resource intensity and prioritization, on the one

¹⁸ O'Mahony et al., 2017a.

Table 6.3
Decision Tree Classification for Different Bandaria Activities

Type	Bandaria Activity	Type of M&E Recommended
SLE	ASCC senior leadership conduct official visit to Bandaria	After-activity reporting (box 8) Path: box 5, no; box 6, no; box 7, no
Exchange	Military Personnel Exchange Program	Impact evaluation (box 14) Path: box 5, yes; box 10, no; box 11, yes Note: The same outcome would occur if activity were seen as a priority rather than resource intensive.
Exercise	Bilateral Bandaria/U.S. Mission Command exercise led by Bandaria Land Forces (BLF)	Performance evaluation (box 19) Path: box 5, yes; box 10, yes; box 15, no; box 16, no; box 17, yes Note: There is strong causal support for many exercises. For the exercise with Bandaria, we assumed there was strong causal support, but that the implementation in Bandaria was not identical to the previous causal evaluations.
Training	BLF pilot/crew members participate in U.S. Army Aviation Center & School training program	Monitoring (box 21) Path: box 5, yes; box 10, yes; box 15, yes; box 20, yes; box 22, yes Note: There is strong causal support for many training programs. For the notional training program with Bandaria, we assumed there was strong causal support, and that the planned implementation, context and goals were similar to previous lessons learned.

SOURCE: RAND analysis.

NOTE: For this table, we assume planners could answer yes for boxes 1 and 2 for all activities.

hand, and how much is known about an activity, on the other. Smaller activities, such as a senior leader engagement (SLE), should be evaluated solely through after-action reporting. In contrast, a recurring resource-intensive activity, such as a military exchange program, is a good candidate for an impact evaluation.

Practical Guidance for Security Cooperation AM&E

Incorporate M&E into ASCC Planning

All activities require some type of M&E, even if it is basic (e.g., AARs), but not all activities require rigorous or intensive M&E. The need for more intensive or rigorous M&E increases as do the need for and benefits from learning and accountability. We created the decision tree to support planners' decisions on the most suitable type of M&E to use when implementing security cooperation activities. At the planning stage, the decision tree will help planners ensure that activities are consistent with higher-level objectives and have valid TOCs. We believe the ASCCs should work with planners to implement the decision tree or a variant that can help guide consistent, high-quality M&E planning.

Consider Costs and Benefits

M&E may be perceived as a large investment for Army security cooperation because it requires coordination across different teams and can become resource intensive. But if it is implemented carefully, M&E benefits should always outweigh costs. ASCCs can achieve the goal

of cost-effective M&E by assessing both the benefits and the costs of M&E at the planning stage. M&E provides value through continuous learning, activity improvement, accountability mechanisms, and early course correction. For instance, M&E would be relevant for cases in which planners are piloting an activity and considering future expansion. M&E provides information about how to improve activity design, correct deficiencies, and improve implementation before further expansion. Costs associated with M&E can be adjusted depending on the type of information and methodological rigor that is required.

Communicate with Units and Partners for Execution

Planners need to communicate with executing units throughout the M&E process. In particular, planners should work with units to understand the type of information required for M&E and how to collect this information, since planners will often rely on executing units for collecting M&E information. For example, task orders already contain some guidance on M&E execution, but many orders do not fully explain the strategic objectives or do not specify what information is needed in M&E templates (e.g., AARs). Planners will also require feedback from units on how to adjust AM&E to ensure planned activities are achievable and useful to executing units. Finally, security cooperation activities are common in conflict-affected or fragile environments, which means there is an increased need for communication with partner nations and intensive in-person collaboration.

Plan for Potential Challenges in M&E Implementation

Security cooperation planners should anticipate future AM&E challenges at the planning stages to the extent possible. Moreover, anticipated risks should be incorporated into the TOC initial assessment, used to adjust the TOC and M&E plan, and taken into account throughout M&E implementation. The following recommendations are designed to help ASCC planners minimize potential challenges:¹⁹

- Ensure there are activity-level TOCs in place that are consistent with the country-level strategy to maximize the chance that activities contribute to strategic goals and have adequate structure and design.
- Incorporate country context into the M&E design and implementation plan, which is especially important for security cooperation activities that take place in unpredictable, fragile, and insecure areas of the world. In addition, data collected on security cooperation engagements may be sensitive or classified, which will require careful planning.
- Collaborate with partner nations that are capable, willing, and flexible enough to conduct M&E and agree on shared goals for security cooperation activities, adapting to partner-specific constraints. For instance, deeply rooted bureaucracy may lead to slow decision-making and implementation, lack of activity flexibility may limit the evaluation design, and lack of capability or willingness to conduct M&E may jeopardize efforts at ensuring accountability.
- Ensure there is a need for the activity before it is conducted. This is built into the TOC initial assessment, but it should be emphasized to minimize the chance that a security

¹⁹ The following bullet points are based on Marie Gaarder and Jeannie Annan, "Impact Evaluation of Conflict Prevention and Peacebuilding Interventions," World Bank Policy Research Working Paper #6496, June 2013; and Dean Karlan and Jacob Appel, *Failing in the Field: What We Can Learn When Field Research Goes Wrong*, Princeton, N.J.: Princeton University Press, 2016.

cooperation activity is later found to have had no positive impact on outcomes because of lack of need (rather than activity ineffectiveness).

- Ensure sufficient technical expertise for M&E. This includes both security cooperation planner capacity and relevant knowledge of or training for units executing activities and collecting M&E data. One option is to develop in-house expertise within the ASCCs on M&E through dedicated teams.
- Adjust for ethical considerations. Security cooperation activities may involve sensitive topics or ethical concerns because of the use of military operations or equipment. Impact evaluations bring additional considerations to help measure security cooperation effectiveness. M&E activities should always take into account the ethical protocols and follow them appropriately.

Task-Level M&E

Before reaching the concluding remarks on the M&E process for security cooperation activities, it is worth emphasizing that while most of this section focused on activity-level AM&E, evaluating security cooperation at the task level remains an important goal for the Army. However, monitoring and evaluating tasks presents a few immediate challenges: first, tasks are difficult to monitor and evaluate because the outputs and outcomes can be driven by exogenous factors outside the scope or control of the tasks, making causal evaluations less reliable. Second, because tasks are typically composed of multiple individual activities, it can be challenging to parse which outcomes are the results of which activities working as expected. The exception to this is when a task is synonymous with one large activity, as with a train and equip task that is composed of one large FMS activity.

Though conducting causal evaluation of tasks is currently out of the scope of this M&E framework, there are other ways that tasks can be evaluated in a meaningful way. ASCC planners can evaluate tasks using causal inference; their ability to do so, however, directly depends on the quality of TOCs. A robust TOC will establish the expected relationships between outputs (and corresponding MOPs), outcomes (and corresponding MOEs), and objectives. ASCC planners' ability to use causal inference is also more possible when task measures are developed with the SMART criteria in mind, and when the evaluation methodology allows for isolating potentially problematic confounding factors—external factors that may also affect the outcomes of interest, and whose effects on the outcomes of interest can be confused with those of tasks.

Greater emphasis should also be placed on collecting qualitative measures for MOPs and MOEs. Qualitative MOPs and MOEs can be collected in a systematic manner, such as using templates to structure subjective assessment. Subjective MOPs and MOEs can provide a robust picture of how well tasks are being designed and executed and complement the causal analysis that is preferable in an activity-level assessment.

Conclusion

This chapter presented a framework to help Army planners determine the effectiveness of security cooperation activities. The framework is complementary with emerging DoD guidance *and* current Army doctrine, and builds on lessons learned from the development community. We identified a set of implementation principles based on three components: initial country

assessments, TOCs, and activity-level M&E. These components serve as generalizable templates that are easy to modify according to context or need. They are meant to serve as living documents that can be modified and updated over time, enabling security cooperation planners to plan for and track security cooperation LOEs. We believe that this framework can enable accountability—for planners, implementers, and stakeholders—and institutionalize a learning process within the Army for the entirety of a project or program's life cycle.

Regardless of whether the Army adopts or adapts this particular framework, we want to highlight a few specific conclusions relevant to any Army security cooperation AM&E process:

- All security cooperation activities require some level of M&E. At a minimum, all activities should have TOCs and AARs. More extensive M&E should reflect planners' need for learning and constraints on effective learning.
- M&E involves costs, but it also provides substantial benefits through learning, activity improvement, and accountability mechanisms; and early correction outweighs their costs. M&E methods should be part of the activity life cycle.
- Security cooperation planners can use monitoring to obtain immediate feedback after an activity takes place and to support future programmatic adjustments. Consistency and interoperability in monitoring allow auditing, reporting, and comparability across different activities.
- Performance evaluation can identify whether security cooperation activities are implemented as planned and whether objectives are being met. Performance evaluations can be conducted in the middle of an activity for midcourse correction or once the activity has been completed to inform future implementation.
- Impact evaluations can help security cooperation planners establish whether and how outcomes are caused solely by an activity and not by something else. Impact evaluations provide the highest-quality evidence of security cooperation effectiveness. Relatively few security cooperation activities warrant large-scale impact evaluations.
- Close communication between security cooperation planners and executing units about M&E plans and implementation is crucial.
- Challenges while implementing M&E are common, and security cooperation planners should plan for these challenges and adjust M&E frameworks accordingly. Planners might consider developing in-house expertise on M&E methods to minimize these potential challenges.

Pilot Tool for Evaluating Security Cooperation Portfolios

AM&E for Army security cooperation activities is key to understanding and maximizing impact. The framework for assessing, monitoring, and evaluating Army security cooperation presented in Chapter Six develops an AM&E process for understanding individual security cooperation *activities*. However, Army security cooperation planners are faced with the challenging task of prioritizing security cooperation activities across strategic objectives—that is, of how to craft effective security cooperation *portfolios*. In particular, there are three main factors that make security cooperation portfolio planning difficult. There is the problem of determining the relative priority of theater objectives while seeking also to balance the requirements of the short-term against the long-term interest. This problem often results not from lax practice but rather genuine trade-offs difficult to resolve definitively across government departments, across services, and among DoD civilian leadership. Further, it is difficult to ascribe direct causal relationships between security cooperation activities and the outcomes they produce. Finally, there is the problem of uncertainty itself. Planners neither possess full information of sufficiently consistent quality on all relevant factors nor can they fully anticipate future trends and occurrences that could profoundly affect security cooperation outcomes.

The problem becomes even more complicated if we recognize that any single security cooperation activity is rarely pursued in isolation from other activities within a partner country. Even within the Army's security cooperation plan for a partner there will usually be a number of activities being conducted simultaneously. As a result, Army security cooperation activities are better thought of as components within an integrated security cooperation portfolio designed to achieve broader strategic objectives.

This chapter reports on a pilot exercise we undertook to explore the possibility that a portfolio-wide view of security cooperation activities could be a useful approach for planners to identify priorities and synergies across security cooperation objectives. We developed a relatively simple proof of concept to examine what a security cooperation planning portfolio analysis tool might entail, and what leverage it could provide planners. We applied the tool to Bandaria, the Army's battle lab, using objectives and activities developed for the Security Cooperation Planners' Course. The tool is designed to provide an integrated view of activity choices and potential outcomes as well as a means for making directly manifest and explicit the different TOCs that would lead to alternative portfolio choices and a means to determine upon what criteria the choice should be made among them to be most consistent with long-term objectives.

With this pilot tool, we believe that it is possible to support a more intense conversation between security cooperation planning and evaluation with results from each being more readily used to inform the other. Explicit consideration of alternative assumptions and their effects

on security cooperation allows performance measurement to tune a TOC such as the simple one we will demonstrate, while the analytical results, in turn, can provide clarity in applying the TOC approach to measurement.

Security Cooperation Portfolio Tool: Inputs and Outputs

The tool was designed in an open-architecture format so that it could provide both transparency and an easily modifiable platform for evaluating outcomes from changing either portfolio inputs or assumptions regarding relationships and therefore outcomes. To serve both these purposes, a model was implemented as an Excel Workbook.

Inputs

For the purpose of this exercise, we considered 20 activity types grouped into five main categories: education and training; talks and conferences; personnel exchanges; knowledge-sharing; and supply, construction, or equipment transfer. These are listed in Table 7.1. We distinguished between activities conducted in the United States and those conducted in the partner nation, bilateral or multilateral activities, short- and long-term efforts, and involvement of civilians, where appropriate.

Table 7.1
Exemplar Security Cooperation Activity Types

		Activity Type	Level of Effort
Education and training	United States	Military only	20
		Military/civilian	0
	Partner nation	Military only	18
		Military/civilian	0
Talks and conferences	United States	Bilateral	1
		Multilateral	1
	Partner nation	Bilateral	1
		Multilateral	0
Exchanges (fixed 1:1)	United States	Long term: U.S. to Partner nation	0
	Partner nation	Long term: Partner nation to U.S.	0
	United States	Short term: U.S. to Partner nation	1
	Partner nation	Short term: Partner nation to U.S.	0
Knowledge-sharing	Partner nation	Emergency response	0
	Partner nation	Multilateral exercise	0
	Partner nation	Joint exercise	270
	Partner nation	Conduct assessment	5
Supply, construct, equip	Partner nation	Provide supply	1
	Partner nation	Build facility	0
	Partner nation	Provide equipment	7
	Partner nation	Conduct R&D	0

SOURCE: RAND analysis.

The user may then enter the level of effort in the right-hand column of Table 7.1 for each activity type in the security cooperation portfolio to be examined. Levels of effort were set in terms of units determined to be most relevant for each of the activity types. The level of effort values included in Table 7.1 are notional values for the baseline portfolio of security cooperation activities we developed for Bandaria. Table 7.2 shows the units in which activity type levels of effort were denominated for this exercise. We include options to measure level of effort in financial or staff costs. For our analysis, we focused on financial costs. At 270 units, the planned joint exercise within the test portfolio showed in Table 7.1 is approximately brigade-sized with some 2,700 U.S. Army participants. The resourcing for each activity type is notional, but reflects values that we observed in our analysis of recent Army security cooperation activities (reported in Chapter Three). As with all input elements in the portfolio tool model, these values may be changed at will to examine the results of differing assumptions or potential surprise.

We modeled the relationship between inputs and outputs in two ways. The first was to introduce a scalar value for each pairing of an activity type with a security cooperation

Table 7.2
Units for Representing Activity Type Level of Effort

Program Costs per Unit of Effort	Unit	Cost^c \$1,000s	Staff Cost^a Days
Education and training: military only (U.S.)	1 person / 3 mo.	35	0
Education and training: military/civilian (U.S.)	1 person / 3 mo.	35	0
Education and training: military only (partner nation)	4 person / 2 mo.	20	0
Education and training: military/civilian (partner nation)	4 person / 2 mo.	15	0
Talks and conferences: bilateral (U.S.)	10 participants	75	0
Talks and conferences: multilateral (U.S.)	10 participants	75	0
Talks and conferences: bilateral (partner nation)	10 participants	40	0
Talks and conferences: multilateral (partner nation)	10 participants	40	0
Exchanges: long term (U.S. to partner nation)	1 person / 6mo.	60	0
Exchanges: long term (partner nation to U.S.)	3 person / 12 mo.	50	0
Exchanges: short term (U.S. to partner nation)	20 participants / 2–6 weeks	80	0
Exchanges: long term (partner nation to U.S.)	20 participants / 2–6 weeks	40	0
Knowledge-sharing: emergency response		^b	0
Knowledge-sharing: multilateral exercise	10 participants	65	0
Knowledge-sharing: joint exercise	10 participants	60	0
Knowledge-sharing: conduct assessment	3 person / 7 days	25	0
Supply, construct, equip: provide supply	3 person / 5 days	60	0
Supply, construct, equip: build facility	5 person / 10 days	80	0
Supply, construct, equip: provide equipment	3 person / 5 days	75	0
Supply, construct, equip: conduct R&D	5 person / 3 mo.	60	0

SOURCE: RAND analysis.

NOTES: ^a There is a column for breaking down costs in terms of dollars as well as actual staff time but the latter exists only as a nonimplemented stub in the current model.

^b Emergency response was considered to be not amenable to prior planning and so was not included in the table.

^c All costs are those borne by the United States and do not include those for the partnering nation.

outcome measure to state the assumed relationship between the two.¹ In other words, how much would one unit of activity type x presumably affect security cooperation outcome y ? In the initial trial we allowed these to vary between -1 and 2 . A value of 0 would imply that a given activity type would be expected to have no effect on the paired security cooperation outcome, while an assessment of -1 would indicate a substantial negative effect and an assessment of 2 a very powerful positive presumed relationship.² How an activity affects each of the Army's multifaceted objectives should match what is included in each activity's TOC on expected outcomes and risks. Table 7.3 shows a sample of these activity type–security cooperation outcome pairings. For example, in our initial analysis, we expected that the effect of multilateral activities on assurance and deterrence would be strongly positive (scale score of 2). In contrast, we thought that the effect of emergency response knowledge-sharing on assurance and deterrence would be minimal (scale score of 0).

The second means for simulating potential effect was also to assign to each such activity type–security cooperation outcome pairing a functional form. These were of four types. The linear form would suggest by its constant slope a direct and immediate correspondence between the scale of the activity type input and the presumed effect on the relevant security cooperation outcome measure. The second form suggested a saturation effect with its constantly decreasing slope; beyond a certain level the system is presumed to receive a diminished benefit from further inputs of that activity type. The third functional form is an S-shaped curve. Initially, there would be relatively little response to an activity type input until a certain critical level is reached, after which the latent system response would kick in for a period of more rapid rise followed, in turn, by a slowing and a saturation level beyond a given point, as was the case with the second form. The fourth functional form, a flat-line response, is redundant with the “ 0 ” presumed effect indicated by the previous set of paired effect values. These functional forms reflect different assumptions about how an activity will contribute to achieving the Army's objectives. These assumptions should match what is included in each activity's TOC.

Outputs

The potential value of the portfolio tool depends on clearly understanding the nature of its outputs—and their limitations. In its present form, it allows security cooperation planners and assessors to implement the TOC concept systematically in an operational setting. Rather than acting as a reliable predictor of outcomes, the tool instead supports full exposition of beliefs regarding security cooperation activities and their implications for outcomes. Put another way, it allows users to draw detailed inferences about what one would need to believe is likely to

¹ For this exercise, we used a set of six idealized outcomes from security cooperation activities with a partner country: assurance and deterrence, improved access and relationships, human and institutional capacity-building, enhancing task and operational capability, increased intelligence and awareness, and joint capability and interoperability. The approach requires that specific statements of causal linkage between activities and outcomes be made even in the absence of supporting empirical or theoretical evidence. While speculative, such statements have the value of rendering the often-implicit notions of causality lying behind recommendations for proposed security cooperation activities open for examination, discussion, and testing of assumptions and uncertainties.

² An example of a negative effect might be associated with an activity providing supplies or equipment. While this might have a positive effect on partner country operational capabilities, one of the six measures of security cooperation outcomes, the literature suggests that it could have a negative effect on another, human and institutional capacities, by increasing the greater potential reward to be gained from corrupt practices. As with all such assessments, these are intended to be modifiable to reflect differing assumptions and perspectives. The goal is to better understand the implications of differing assumptions for alternative portfolios and their outcomes rather than serving as hardwired, fixed assessments.

Table 7.3
Scale Effect Assumptions for Assurance and Deterrence Objective

	Activity Type	Scale
Effect on assurance and deterrence	Education and training: military only (United States)	1
	Education and training: military/civilian (United States)	1
	Education and training: military only (partner nation)	1
	Education and training: military/civilian (partner nation)	1
	Talks and conferences: bilateral (United States)	1
	Talks and conferences: multilateral (United States)	1
	Talks and conferences: bilateral (partner nation)	1
	Talks and conferences: multilateral (partner nation)	1
	Exchanges: long term (United States to partner nation)	1
	Exchanges: long term (partner nation to United States)	0
	Exchanges: short term (United States to partner nation)	1
	Exchanges: long term (partner nation to United States)	0
	Knowledge-sharing: emergency response	0
	Knowledge-sharing: multilateral exercise	2
	Knowledge-sharing: joint exercise	2
	Knowledge-sharing: conduct assessment	0
	Supply, construct, equip: provide supply	1
	Supply, construct, equip: build facility	1
	Supply, construct, equip: provide equipment	2
	Supply, construct, equip: conduct R&D	0

SOURCE: RAND analysis.

be true in order to advocate one portfolio of activities over another in the pursuit of security cooperation objectives. Further, its strength lies not in providing individual “forecasts” but in allowing security cooperation planners and evaluators to set up compound computational experiments and then assisting them in reasoning over the ensembles of results generated by changing assumptions regarding either the appropriate level and mixture of activities within a country portfolio or their presumed effects. This point will be demonstrated in the following example.

In its current format the tool provides outputs of the character shown in Table 7.4.³ These results are obtained by running the portfolio of activities shown in Table 7.1 against a specific

³ Several other capabilities are present within the current Excel model or have had currently nonimplemented stubs provided for them for later development and elaboration. These will not be detailed more fully at this time. Although the costs are reported in dollars, the potential effect scores are meaningful only in comparison with each other—they do not correspond to a real-world unit of account.

Table 7.4
Presumed Costs and Potential Effect of Portfolio

Outputs		
Costs	Program Monetary Costs (\$1,000s)	18,240
Potential effects on	Assurance and deterrence	651
	Access and relationships	405
	Human and institutional capacity	368
	Task and operational capability	477
	Intelligence and awareness	543
	Joint capability and interoperability	635

SOURCE: RAND analysis.

set of assumptions (such as those shown, in part, in Table 7.3) and then aggregating the presumed response for each of the six categories of security cooperation objectives.⁴

Sample Portfolio Analysis for Bandaria

We have used the treatment of the fictional country of Bandaria⁵ as the basis for illustrating uses for a portfolio tool based on the principles that we applied in the Excel portfolio model. Bandaria was developed for the Army Security Cooperation Course (ASCPC) to teach class participants how to think through planning security cooperation activities and the country task assessment process. Students are provided with characteristics and capabilities for a partner nation for which the U.S. Army is trying to develop its relationship as well as to enhance regional partner capabilities. The ASCPC materials provide detailed information on Bandaria, including the GCC strategy for Bandaria and its neighbors, the ASCC tasks, and timelines for building capability. Students are then expected to develop country plans through a five-step process: (1) mission analysis, (2) plan development, (3) assessment plan, (4) activities plan, and (5) resource plan.⁶ In this section we address a series of questions that planners should take into account as they develop security cooperation portfolios.

What is the presumed effect of the current portfolio of activities for the Bandaria security cooperation program?

We have already shown one portfolio of activities applied to one country, the fictional Bandaria, and their presumed collective effect on security cooperation objectives in Tables 7.1

⁴ The implicit assumption in this formulation is that there are no indirect interactions between LOEs and that the results from each specific LOE and its scale of effort can be aggregated with the outcomes from the other LOEs in the same portfolio for the same security cooperation outcome measure. We have also abstracted the period during which outcomes from LOEs enacted in a particular year may be expected to be realized. The potential for charging more complex interactions among LOEs, as well as lags between LOE performance and security cooperation outcome, is latent within the portfolio tool but was not implemented for this demonstration.

⁵ This example was in turn constructed from the material found in Department of the Army, 2016.

⁶ Department of the Army, 2016, p. 54.

and 7.4, respectively. The outcomes reflect the TOC embodied in the scale effect and functional form relationships presumed for each pair-wise association between each activity type and each security cooperation objective. A different TOC and a different prioritization of importance among security cooperation objectives would probably lead to an alternative formulation of the portfolio of LOE activities for Bandaria for the year in question.

How well might other potential portfolios of similar cost perform?

We placed the Excel model within the Computer Assisted Reasoning® system (CARs™) software environment to carry out structured compound computational experiments. Using CARs we generated a set of 2,000 randomly assembled alternative security cooperation portfolios for Bandaria by setting ranges for each of the potential activity types and drawing results from this range.⁷ Within this set of alternative portfolios, we discovered a set of 27 for which total cost was either equal to that for the current Bandaria security cooperation portfolio or slightly less. As an exercise in determining the effect of alternative TOC assumptions, we initially evaluated their performance by first presuming that all activity types had a uniform scale effect of 1 on each of the six security cooperation objectives. The resulting performance of these alternatives under the assumptions making up that implicit TOC is shown as the colored lines in Figure 7.1. These may be compared with the base case or the current Bandaria portfolio's performance (the values from Table 7.1), shown as a thick black line. The six security cooperation outcome measures are arrayed along the horizontal axis with each line showing the relative performance of the associated alternative security cooperation portfolio across all six measures.

Are there any alternatives that should be examined in detail in preference to the current program?

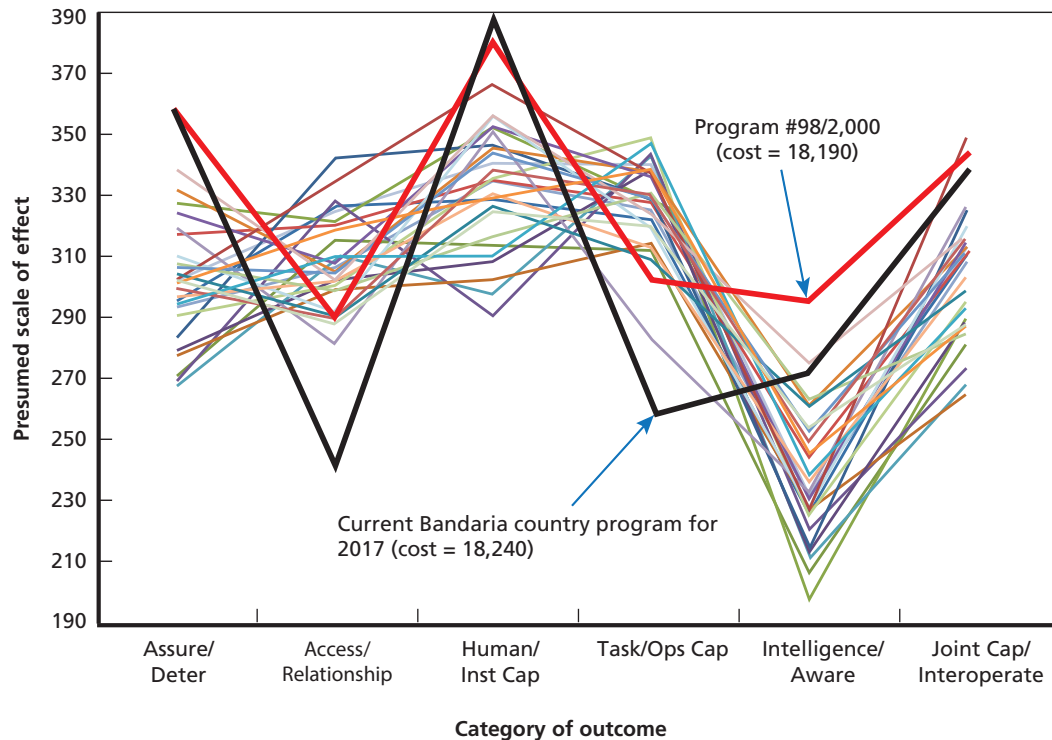
Alternative portfolio #98 out of the set of 2,000 random portfolios appears to have properties worth understanding. It is marginally cheaper than the base case Bandaria portfolio of security cooperation actions, and while it resembles the latter's presumed effect on three of the security cooperation objectives, it has a clearly superior presumed performance for the three others.

What are the differences in TOC implied by the current Bandaria and the candidate alternative security cooperation portfolio?

Table 7.5 illustrates that this portfolio implies a significantly different TOC from that underlying the current Bandaria portfolio. Whereas the current plan implies that a main driver to achieve successful security cooperation outcomes is to emphasize educational programs, both inside the United States and in Bandaria itself, as well as to hold a major brigade-sized joint exercise, the TOC implicit in plan #98 is to scale back joint exercises and to shift emphasis away from educational activities (and pitch those that remain much more toward U.S.-based instruction rather than in Bandaria itself) and toward those involving exchanges, activities virtually disregarded by the current Bandaria plan.

⁷ For this purpose we implemented a Latin Hypercube design so that the ensemble of 2,000 alternative portfolios sampled evenly across the ranges of possible values set for scale of each LOE effort. Those LOEs that had been set to zero in the current Bandaria security cooperation portfolio were also set to zero in the alternative portfolios of LOE actions.

Figure 7.1
First Comparison of Security Cooperation Portfolios



SOURCE: RAND analysis.

NOTE: Presumed effect of current "Bandaria" program compared to 27 random portfolios of LOEs across six outcome measures (all effect assumptions equivalent) with total program cost between 18,000 and 18,240.

RAND RR2165A-7.1

The presumed effects of these changes in emphasis and therefore the portfolio of activities implied by different underlying TOCs are shown in Figure 7.2.

What if we apply different assumptions about the scale, direction, and functional form of effects—a radically different implicit TOC? How will that affect the inferences we draw about program effect?

In the prior analysis we presumed a uniform effect across activity types on each security cooperation objective.⁸ We now provide specific scale factors between -1 and +2 for each activity type/security cooperation objective pairing based on an assessment of the general relationship between activity types and intended (as well as potential unintended) outcomes and the particularities of the Bandaria case. One might view these two separate evaluations as an examination of the consequences of applying two different TOCs to the same combined sets of activities. How would outcomes be affected if we believed in one of the TOCs but in reality disclosed

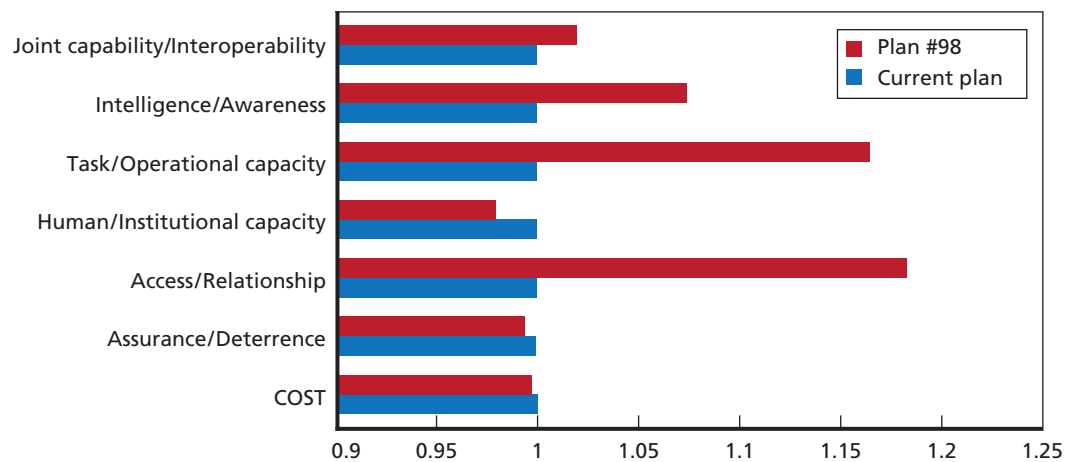
⁸ This was represented by initially setting each of the values in Table 7.3 to 1 (along with the corresponding scale effect inputs for the other five security cooperation outcomes) rather than the values shown in that table. Those results are the ones that have been reported to this point. For the balance of the experiment we shifted to the individually assessed values shown in Table 7.3 as well as those for the corresponding pairings of each LOE with the other five security cooperation outcome measures.

Table 7.5
Comparison of Components of Two Alternative Portfolios

Inputs	Current Plan	Plan #98
Education and training: military only (United States)	20	5
Education and training: military only (partner nation)	18	1
Talks and conferences: bilateral (United States)	1	8
Talks and conferences: multilateral (United States)	1	1
Talks and conferences: bilateral (partner nation)	1	14
Exchanges: long term (United States to partner nation)	0	10
Exchanges: long term (partner nation to United States)	0	10
Exchanges: short term (United States to partner nation)	0	3
Exchanges: long term (partner nation to United States)	1	4
Knowledge-sharing: joint exercise	270	219
Knowledge-sharing: conduct assessment	5	13
Supply, construct, equip: provide supply	1	22
Supply, construct, equip: build facility	0	5
Supply, construct, equip: provide equipment	7	1

SOURCE: RAND analysis.

Figure 7.2
Comparison of Costs and Presumed Outcomes of Two Alternative Portfolios



SOURCE: RAND analysis.

NOTE: The values of results for the current security cooperation plan for Bandaria have all been set to 1.0 for comparison and then plan #98 is scaled accordingly.

RAND RR2165A-7.2

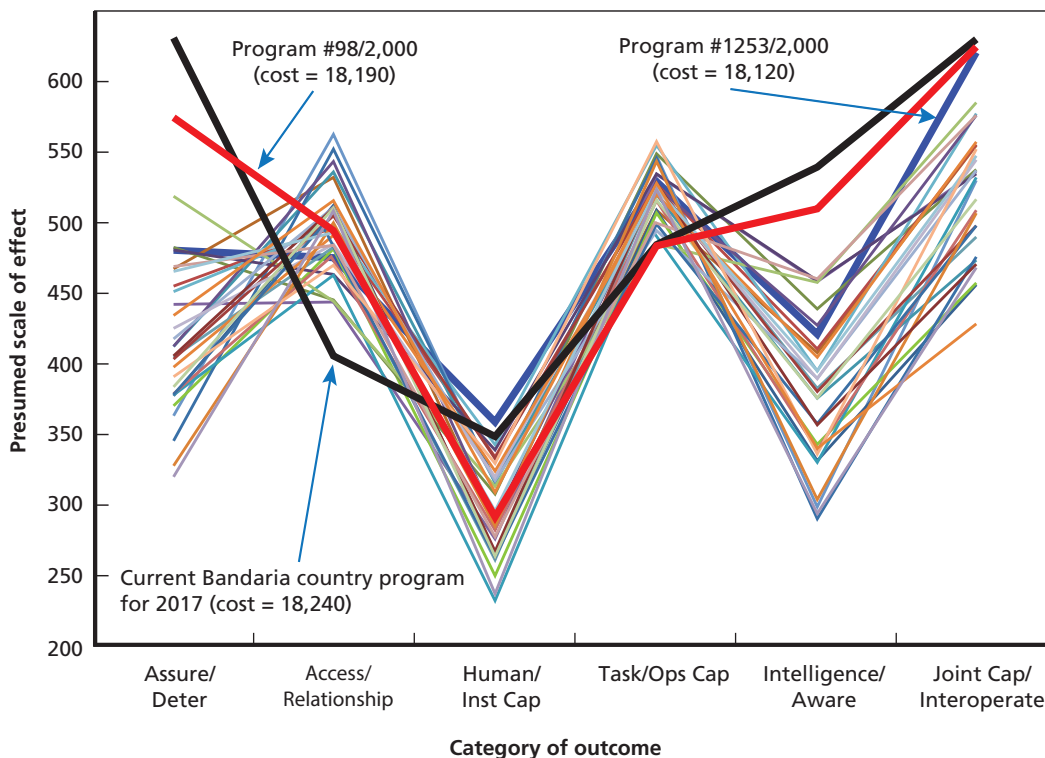
some reason to question its underlying assumptions? In this manner, putting the TOC into these explicit terms as part of a performance measurement protocol could help implement a “reconnaissance by fire” approach to making the most of the feedback from evaluation to modify or enhance confidence in an underlying TOC.

Figure 7.3 shows the result of this experiment. It suggests that the confidence placed in the current Bandaria portfolio of security cooperation activities may be more solidly based than the prior, less sophisticated exercise in Figure 7.1 might have led us to believe. The current program does relatively poorly in the security cooperation objectives of access/relationship and task/operation capacity, but while the candidate alternative, portfolio #98, may be expected to perform considerably better in the former, for the other five objectives it does no better than, and in several cases worse than, the current portfolio of activities planned for Bandaria. Instead, under the TOC implied by the more nuanced set of scaling factors and functional forms now being used in the analysis, another program, #1253, now appears to be one that might usefully draw the attention of the security cooperation planners.

Can we discover something systematic about the instances of relative poor performance of the current Bandaria security cooperation plan that we should understand to help us improve it?

We also wanted to test the vulnerability of one portfolio to changes in the assumptions underlying it. Whereas we just generated a set of 2,000 potential alternative security coopera-

Figure 7.3
Second Comparison of Security Cooperation Portfolios



SOURCE: RAND analysis.

NOTE: Presumed effect of current “Bandaria” program compared to 30 random portfolios of LOEs across six measures of outcome (effect assumptions varied by LOE and outcome measure) with total program cost between 18,000 and 18,240.

RAND RR2165A-7.3

tion portfolios for Bandaria above to examine the relative strength of each under consistent assumptions, we now use the tool to help us understand what leads to success for a specific portfolio of activities. Keeping the current Bandaria program fixed, we now instead generate 20,000 random combinations of scale effect assumptions that are each allowed to vary within limits set for them.⁹ We then examined the performance of the current Bandaria program under the assumptions represented by each of these cases. We set as a success criterion that the current Bandaria program outcome could not be viewed as favorable under a given case unless it attained at least the average level of performance for all six of the security cooperation objective categories.¹⁰ We then looked across the entire resulting dataset to determine what was systematic in setting apart successful from unsuccessful outcomes for the Bandaria security cooperation plan.

Using data mining techniques, we were able to conduct an automated search to find lower-dimensional regions that could best explain what led to the different outcomes.¹¹ (We call these “boxes,” defined by discovering very small subsets of the full 84 dimensions across which we search, which are sufficient to describe the region where cases of interest are clustered.) The software first proposes the “peeling trajectory” shown in Figure 7.4. It shows a series of possible boxes, defined by identifying specific regions along several of the axes of uncertainty that are designed to describe the leading common factors among the cases of interest.

Figure 7.4 conveys how there is a trade-off between coverage (the share of all cases of interest falling within the parameters suggested for a given proposed box) and density (the share of cases falling within the proposed box that are truly cases of interest and not false positives). A third factor that may lead to the selection of one box over another is how intelligible the scenario conditions implied by the proposed parameters might be to those doing the planning or implementation of the actual course of action to be followed. The numbered points in Figure 7.4, lying in the region of closest correspondence between the two desirable properties of coverage and density, have been selected because they are frequently of most interest to planners.

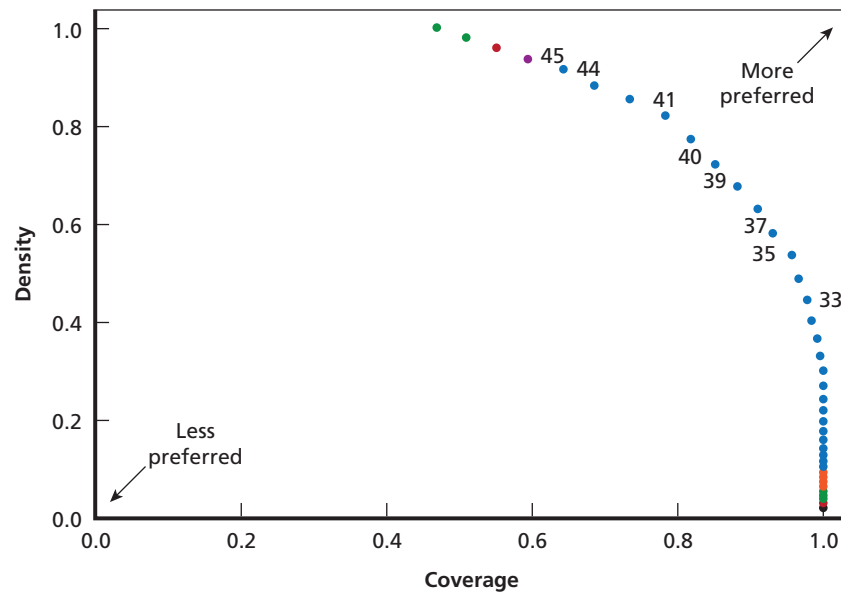
Table 7.6 shows the coverage and density characteristics for each of the proposed boxes associated with the numbered points in Figure 7.4. All but one propose a dimensional collapse to points along only six dimensions of the total 84 as being sufficient to describe the conditions of each box. Table 7.7 provides detail on one example, the box associated with point 39 in the peeling trajectory, which contains 85 percent of all the cases of interest within the full sample

⁹ This involved generating 20,000 combinations of alternative settings for each of the 84 scale factors governing the effectiveness of the subset of activities used in this analysis. (The full model contains 120 such scaling factors as well as 120 functional form factors.) Once again, a Latin Hypercube experimental design was utilized to ensure uniform sampling throughout this highly dimensional space.

¹⁰ For each of the six outputs, all 20,000 results were scaled by setting the highest value attained at 1 and scaled all the others proportionately. Therefore, a case needed to score at least 0.5 for each of the six security cooperation objectives for the case as a whole to be deemed successful.

¹¹ These boxes may be thought of as analytically derived scenarios of interest. In this case, they describe in precise terms what TOC would need to hold true for the observed favorable results to be favored or expected. This analysis used an R tool called “sdtoolkit” (i.e., scenario discovery toolkit), which implements the patient rule induction algorithm. Jerome H. Friedman and Nicholas I. Fisher, “Bump Hunting in High-Dimensional Data,” *Statistics and Computing*, Vol. 9, No. 2, 1999; Benjamin P. Bryant and Robert J. Lempert, “Thinking Inside the Box: A Participatory Computer-Assisted Approach to Scenario Discovery,” *Technological Forecasting and Social Change*, Vol. 77, No. 1, January 2010. The required calculations for a database consisting of 84 input variables, each of which vary in a set of 20,000 total cases, takes several minutes to run on a high-performance server.

Figure 7.4
Peeling Trajectory of Alternative Specifications



SOURCE: RAND analysis.

NOTE: Peeling trajectory of alternative specifications of conditions leading to higher consistent levels of success.

RAND RR2165A-7.4

Table 7.6
Density and Coverage Measures for Numbered Boxes in Figure 7.4

Box Index	Density	Coverage	Dimensions
33	0.44	0.98	6
35	0.54	0.96	6
37	0.63	0.91	6
39	0.72	0.85	6
40	0.77	0.82	6
41	0.82	0.78	6
44	0.91	0.64	6
45	0.93	0.59	7

SOURCE: RAND analysis.

(that is, those meeting the criterion for being judged a successful outcome), with 28 percent of the cases in the box not meeting the criterion.

The conditions shown in Table 7.7, along with the other numbered boxes in the same region, tell a consistent story that provides insight into both the Bandaria plan's performance and the TOC that informs it. Of all the factors involved, the crucial element is the scaling relationship between the conduct of a joint training exercise in a partner country (KPJX) and the potential effect on each of the six areas of security cooperation performance. The Bandaria plan is notable for the preponderance of this activity within the portfolio as well as the

Table 7.7
Conditions Associated with Box 39 in Figure 7.4

Scaling Factor for Joint Exercise	Relationship	Bound
Intelligence/Awareness	>	0.9980
Operational capacity	>	1.0168
Joint capability/Interoperability	>	0.9155
Human/Institutional capacity	>	0.9349
Assurance/Deterrence	>	0.8636
Access/Relationship	>	0.8717

SOURCE: RAND analysis.

scaling factors associated with this activity type in the implicit TOC rendered explicit by the analysis shown for this example. Whereas the scaling factors shown in Table 7.7 were all set at 1 (as were all others) for the experimental results shown in Figure 7.1, the same factors were mostly set at the extreme value of 2 for the analytical results shown in Figure 7.3 (as well as Figure 7.5).¹² The Bandaria plan places great emphasis on joint exercises because of the relatively large return on this investment presumed by the TOC.

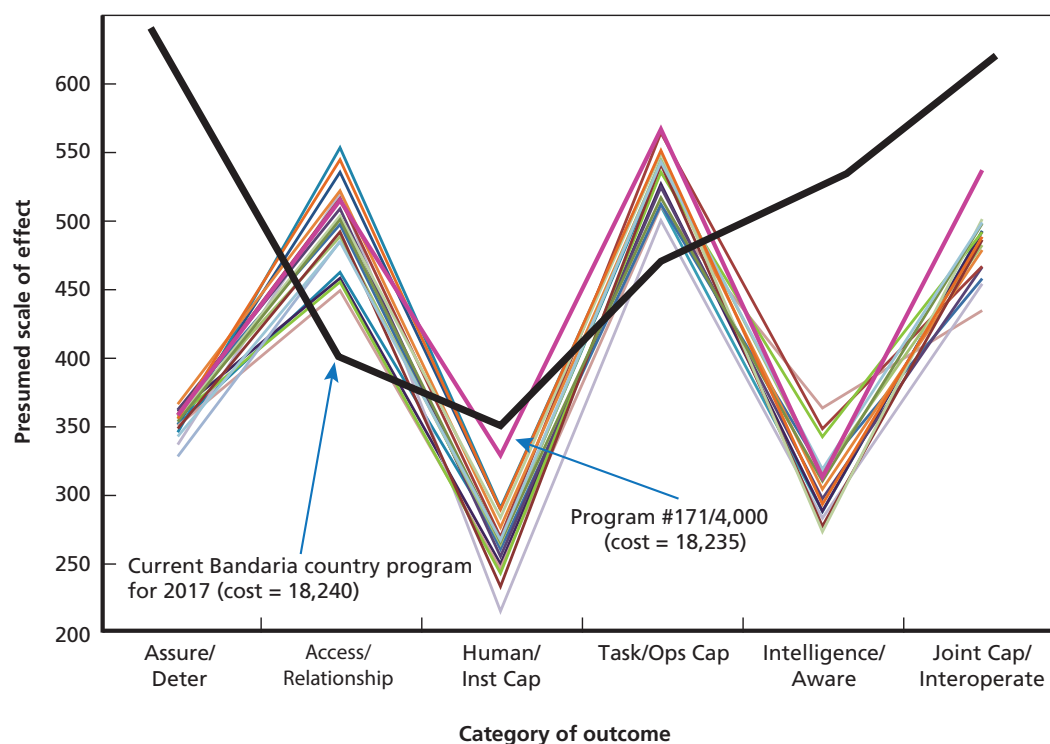
This consideration explains the potential effect of both in-country joint exercises and the current Bandaria plan when compared with the other alternative portfolios we have examined. It also illuminates how a modeling infrastructure containing several hundred variables (and thus degrees of freedom) may be approached to determine which among them are the key drivers of outcomes. This renders the TOC exposition and analysis process feasible, tasks that would be almost impossible to perform without such tooling support. One could now examine alternative specifications of the TOC by manipulating the scaling factors (as well as those mandating the functional forms, a further 84 specifications of inputs), but the results in Table 7.7 suggest that this TOC might be fairly robust to change. Even if all of the scaling factors were set to 1, thus halving the effect of joint exercises, the cases judged to be successful would most likely still be found in the region described by the parameters of box 39.

Does this mean that only by conducting brigade-sized joint exercises we will achieve the desired results? What if we fall short? Can we discover alternative portfolios that would still be similarly successful if forced to cut back?

We return once more to search through portfolio space while maintaining the same scaling effects and form parameters that define the operative TOC presumed for the current Bandaria plan. In this instance, we randomly generated 4,000 alternative portfolios for Bandaria by the same means we used previously. The only constant was that the level for the in-country joint exercise was trimmed to a consistent one-third of the level mandated in the current Bandaria plan for 2017, effectively a joint exercise with an augmented battalion rather than a full brigade. The results are shown in Figure 7.5.

¹² The presumed scaling effect on the human and institutional capacity-building objective for joint exercises in-country was set at the value 1 in this conceptualization of the operative TOC.

Figure 7.5
Third Comparison of Security Cooperation Portfolios



SOURCE: RAND analysis.

NOTE: Presumed effect of current "Bandaria" program compared to 23 random portfolios of LOEs (joint exercise cut back from brigade to augmented battalion) across six measures of outcome (effect assumptions varied by LOE and outcome measure) with total program cost between 18,000 and 18,240.

RAND RR2165A-7.5

Once more the full set of generated portfolios was winnowed to yield 23 candidates with a total cost close to but not exceeding that for the current Bandaria plan. These plans can come close to or even exceed the current plan in some of the output categories but would suffer from the cutback if the prime interest lay in enhancing assurance and deterrence or joint capability and interoperability. Under the prevailing TOC, the cutback from brigade to battalion in conducting a joint exercise with Bandaria would come at some cost. However, having enumerated these potential losses in effectiveness while maintaining the same set of assumptions, planners may now discuss among themselves, their superiors, and other U.S. federal-level actors what consequences might be presumed and what further adjustments may be possible.

We can focus on portfolio #171 among this new group of 4,000 generated alternatives to the current plan. It can best be understood as a candidate that places different emphases on security cooperation outcomes in the face of the mandated cutback rather than one providing a clearly superior allocation of resources in all respects as we had sought before. Table 7.8 displays the nature of the trade-offs proposed among activity types and how the new allocation of resources compares with that for the current Bandaria security cooperation plan. Keep in mind, however, that the presumed effects in aggregate will very much depend on the TOC enunciated in the form of the various factors that translate LOE levels into presumed effects

Table 7.8
Comparison of Two Alternative Portfolios with Joint Exercise at Different Scales

Inputs	Current Plan	Plan #171
Education and training: military only (partner nation)	20	21
Education and training: military only (partner nation)	18	29
Talks and conferences: bilateral (United States)	1	8
Talks and conferences: multilateral (United States)	1	16
Talks and conferences: bilateral (partner nation)	1	21
Exchanges: long term (United States to partner nation)	0	13
Exchanges: long term (partner nation to United States)	0	30
Exchanges: short term (United States to partner nation)	0	23
Exchanges: long term (partner nation to United States)	1	25
Knowledge-sharing: joint exercise	270	90
Knowledge-sharing: conduct assessment	5	14
Supply, construct, equip: provide supply	1	28
Supply, construct, equip: build facility	0	16
Supply, construct, equip: provide equipment	7	6

SOURCE: RAND analysis.

on security cooperation objectives. If, for example, the cutback in the scale of joint exercise was based on a revised conception for the underlying TOC, perhaps based on the results gained through more detailed and systematic assessment and evaluation, then the consequences of the radical shift in emphasis toward other security cooperation activities might prove less consequential for those two objectives for which the current Bandaria plan appears to be the most effective under the prior TOC.

Conclusion

The interplay between TOC and analysis presented in this chapter illustrates the value of security cooperation portfolio planning, and how the activity-level AM&E process outlined in Chapter Six works within broader strategic planning needs. The tool and method described here are not intended to provide a *deus ex machina*, an algorithmic system that will obviate the role of security cooperation planners. Quite the reverse: it aids planners in using more rigorous analysis.

The portfolio tool and the robust decisions approach may accomplish several goals for security cooperation planners and evaluators:

- render explicit **underlying assumptions** regarding the relationship between security cooperation activities and desired goals and outcomes in a way that is made concrete for service member use and engagement

- provide a **venue for discussion** of a type supporting the generation of results such as those shown in Appendix D that is valuable but otherwise difficult to coalesce in practice
- **tie planning more organically to evaluation** by providing a framework for incorporating evaluative insights operationally
- specify and **address uncertainty** by testing exposure to various alternative assumptions
- discover **new courses of action** or sequencing of actions
- develop a **sophisticated appreciation within a complex space** of alternative courses of action, their potential consequences, and the basis on which to make choices among them consistent with long-term security cooperation objectives
- conduct analysis in the **absence of an explicit TOC** and for which the TOC is itself among the outputs of the analysis.

The goal is to support a more analysis-based dialogue between planners and evaluators with results from each being more readily used to inform the other. Results from performance measurement may be used to tune a planning model such as the simple one we have constructed, while the model results, in turn, can provide clarity in applying the TOC approach to measurement.

Conclusion

As security cooperation continues to be an important instrument for the Army to achieve its objectives, the Army needs a better understanding of relative effectiveness. Part I of this report found that recent Army security cooperation activities have aligned fairly well with what previous analyses have found contribute to effectiveness in both security cooperation and international development assistance. However, the lack of systematic AM&E across Army security cooperation activities has made it impossible to examine individual and aggregated activities in a way that truly drives prioritization efforts. This reflects both the inherent challenges that exist in evaluating security cooperation and the lack of robust AM&E processes throughout the Army and DoD more broadly.

OSD has worked to improve DoD enterprise-wide security cooperation AM&E procedures. Its recently released DoDI has enumerated requirements that are in keeping with good practices adopted by international development organizations and recognize the complexity of U.S. security cooperation in toto.¹ However, OSD's focus has been at the policy level and its guidelines serve primarily to provide guidance for GCC planners.

To help the Army better plan and execute its security cooperation activities, we developed AM&E implementation guidelines based on emerging OSD guidance, current Army processes, and good practices adopted by international development organizations. We expect the primary users of this framework and portfolio tool will be ASCC planners, while HQDA staff can use it in their oversight role. ASCC planners serve as the primary interface between GCCs, the department of the Army, and Army implementers. We believe that adopting the Army security cooperation AM&E framework presented in this report will provide a good foundation for deepening the Army's understanding of its security cooperation effectiveness.

The Army's focus on deepening its security cooperation AM&E procedures at the activity level can serve as a model for how DoD links its strategic guidelines to operational planning. The Army's bottom-up operational approach can be a valuable complement for OSD's top-down strategic approach, particularly with the help of the framework and portfolio tool developed in this report. Taken together, the OSD and Army approaches form the basis for a comprehensive process that accounts for strategic, operational, and tactical considerations in support of decisionmaking.

¹ Office of the Secretary of Defense, 2016.

What Is Security Cooperation?

This appendix provides information about security cooperation activities and objectives. Understanding what security cooperation entails is important for assessing how to improve AM&E for Army security cooperation.

Security Cooperation Activities

Security cooperation encompasses a diverse set of partner engagement activities, ranging from small-scale exchanges of experts or consultations among high-ranking officers to large-scale combined military exercises. From the G-TSCMIS dataset discussed in Chapter Three, we identified ten categories encompassing the most commonly used security cooperation activities. The categories are not intended to be exhaustive, but rather are meant as illustrations of some of the choices the Army makes on how to spend finite resources conducting security cooperation activities in pursuit of strategic objectives.

Assessments

Army-conducted assessments represent a small portion of the yearly events undertaken, and are most often conducted with nascent or developing partners that wish to build capacity in certain areas. Common assessments include an Army Corps of Engineer survey of water and health sanitation infrastructure. Assessments are typically followed by other types of activities, including medical trainings, or, in the case of an assessment oriented toward equipment acquisition, followed by FMF/FMS.

Conferences

Conferences occur with great frequency, but at a low per-event cost, within the annual Army security cooperation portfolio. All GCCs hold conferences at similarly high rates with the exception of PACOM and SOUTHCOM, where conferences are more frequent. Conferences are commonly used for information-sharing and agenda-setting within a GCC.

Senior Leader Engagements

Senior leader engagements (SLEs), where senior representatives of the U.S. Army meet with their counterparts for information exchanges, short trainings, and other relationship-building activities, represent a relatively small portion of annual security cooperation events. However, the Center for Army Lessons Learned publication *Security Cooperation: Lessons and Best Practices* of 2016 contends that SLEs are a cost-effective priority activity for the Army, based on the

low-cost opportunity to meet partners.¹ SLEs are a particularly common tool in the security cooperation portfolio for AFRICOM and NORTHCOM, which may speak to their use for regular communication between established allies and nascent partners.

Staff Talks

The Army promotes bilateral staff talks as key to developing professional partnerships and encouraging interaction between partner forces. Staff talks typically involve a small number of senior staff meeting counterparts to discuss specific topics. In SOUTHCOM, staff talks have been institutionalized as annual events with partners, whereas in other GCCs they occur on an ad hoc basis.² Staff talks also occur frequently in AFRICOM, where they have been supported for several years by the African Contingency Operations Training Assistance program, and, more recently, representatives from the African Union have conducted multilateral staff talks with AFRICOM.

Subject Matter Expert Exchanges

Like SLEs, subject matter expert exchanges (SMEEs) are typically short interactions with a limited number of personnel on the United States and partner nation sides. Unlike SLEs, SMEEs typically occur on the request of the partner nation for a specific purpose, for example, requesting a SMEE on Army logistics and personnel management for a weeklong course or training. The Army does the highest number of SMEEs annually with NORTHCOM and EUCOM, where high-capacity partners request technical SMEEs on issues like cyber defense and chemical-biological weapons defense.

Humanitarian and Civic Assistance

Humanitarian and civic assistance is a significant slice of the annual Army security cooperation activities and budget. As a large and heterogeneous category, humanitarian and civic assistance activities include short medical and veterinarian civic action programs, crisis management courses, and civilian-military coordination trainings. While the majority of humanitarian and civic assistance activities are not incredibly resource intensive, their frequency means that, taken together, they represent a significant portion of the cost of Army security cooperation activities. Encompassed within the humanitarian and civic assistance category are also emergency and disaster response activities, which occur infrequently but are resource intensive when they do happen. The Army conducts civilian assistance programs in all GCCs, but those programs vary widely across the GCCs.

Exchanges

As a portion of total Army security cooperation activities, personnel exchanges are very common (averaging over 200 instances per year) but relatively inexpensive. Exchanges vary in length and may be as short as four to six weeks or as long as two to three years, in the case of the Military Personnel Exchange Program, which supports a limited number of personnel every year to participate in exchanges with similar personnel in partner countries for a two- to three-year period to encourage collaboration, create personal relationships between U.S. and

¹ Center for Army Lessons Learned, 2016, p. 15.

² Lt. Col. Antwan C. Williams, "U.S. Army South Conducts Army-to-Army Staff Talks with Chile," *U.S. Army South Public Affairs*, October 31, 2011.

foreign personnel, and enhance strategic partnerships.³ The Army conducts the majority of exchanges with NORTHCOM and AFRICOM, reflecting the use of exchanges to enhance interoperability between closer partners as well as building a greater level of trust and mutual understanding with newer partners.

FMF/FMS

As a set of programs, FMF/FMS is intended to create more capable partners, while also promoting higher capacity of the personnel responsible for using and maintaining U.S.-provided equipment and supplies.⁴ While FMF/FMS activities are fairly evenly spread across most GCCs, CENTCOM stands out for receiving the highest percentage of equipment and supplies overall.

Training

Training encompasses a wide set of programs with a number of different objectives, personnel requirements, and implementation timelines. Training is generally held to provide a partner with a specific capacity, as with counter-improvised explosive device training and chemical, biological, radiological, and nuclear countermeasure training, or to enhance the professionalization of a partner, as with a leadership training course. The type and size of the training activity can vary greatly, from a small team (two or three individuals) conducting an initial survey of a partner's current capacity and capabilities for a limited number of days, to large-scale development of a training program and associated infrastructure to facilitate institutionalization of equipment, supplies, and training in a partner's military. Programs like International Military Education & Training courses, which are offered to both military and civilian government officials in partner nations, also teach foundational elements of U.S. military doctrine including civilian-military relations, civilian control of the military, international humanitarian law, and other topics designated by the U.S. military to be critical for creating professional military forces.⁵

The Army does the highest number of trainings with PACOM, NORTHCOM, and AFRICOM, though the wide range in structure, personnel requirements, and duration across trainings makes creating a unified definition of the category challenging.

Exercises

Exercises, like exchanges, are some of the more common security cooperation activities that the Army engages in. The majority of exercises are small, tabletop exercises with fewer than 30 participants, lasting for approximately a week, to plan for and/or simulate a war-fighting scenario. However, a smaller number of exercises are large, battalion-size field exercises requiring a significant input of resources and personnel, and are therefore significantly more expensive. As a subset of exercises in total, humanitarian exercises are less frequent but more expensive than the average exercise because they typically require the coordination of multiple agencies with the U.S. government as well as partner nation government agencies, and potentially

³ MAJ Donald R. Meeks, Jr., and MAJ Chad Froehlich, "What Is MPEP?" *Military Police*, Vol. 19-08-2, Fall 2008.

⁴ Defense Security Cooperation Agency, "Foreign Military Financing (FMF)," n.d.a.; Defense Security Cooperation Agency, "Foreign Military Sales (FMS)," n.d.b.

⁵ Defense Security Cooperation Agency, "International Military Education and Training (IMET)," n.d.c.

international organizations and nongovernmental organizations (NGOs). The Army uses exercises across all GCCs to enhance interoperability, conduct scenario planning, and build partner capacity in specific areas. In PACOM, AFRICOM, and NORTHCOM, the Army conducts major bilateral and multilateral exercises on an annual basis.⁶

Security Cooperation Spectrums of Engagement

When considering the activity portfolios of different countries, specific types of activities tend to co-occur with each other because they collectively address the common security challenges that specific types of countries tend to face. For ongoing RAND research examining the Global Landpower Network concept,⁷ we assessed security cooperation activities included in G-TSCMIS, 2009–2014. From the event and partner country characteristics, we identified three key spectrums of activity that tend to correspond to specific portfolios of activities:

- Shaping the theater—activities to anticipate and ameliorate potential security challenges
- Building partner military capabilities—activities that make partner militaries more able to ensure domestic security and be a stabilizing force internationally
- Deepening partner compatibility—activities that improve the compatibility of partners with the United States, both politically and militarily.

Each of these spectrums can be further divided into halves, depending on the quality of the U.S.-partner relationship required. This yields six distinct categories of activities. Figure A.1 depicts these types.

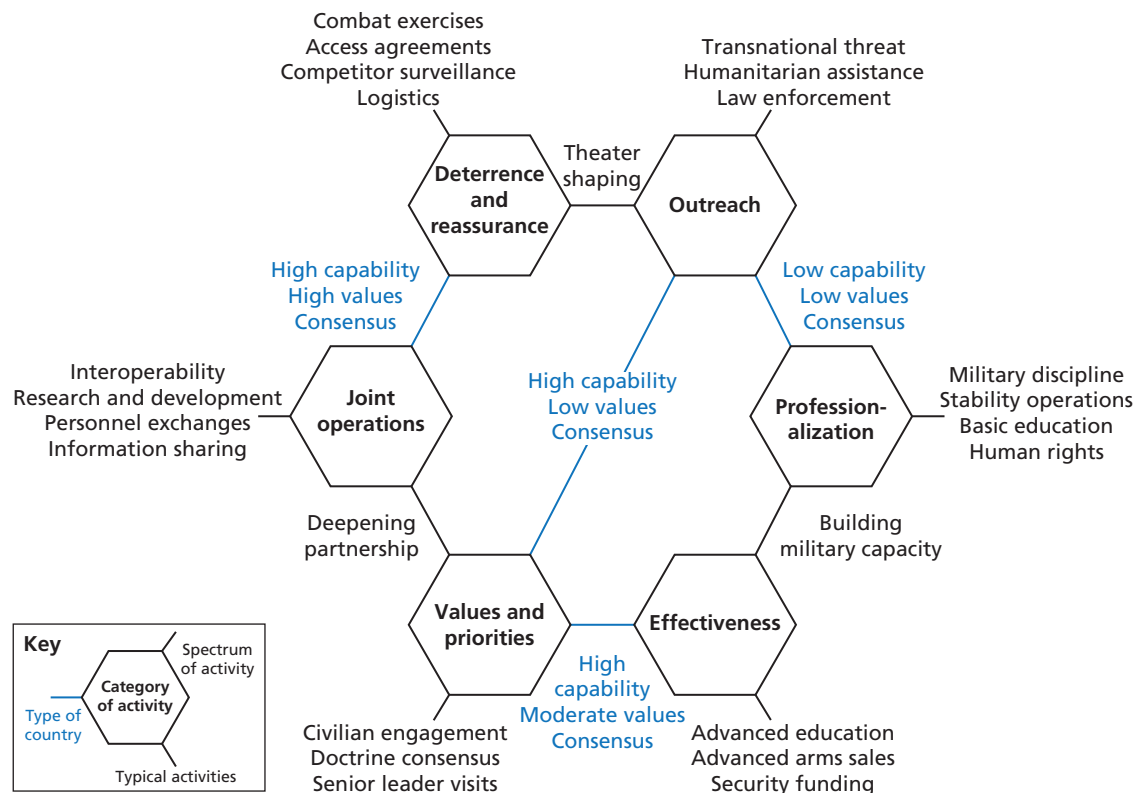
Deterrence and reassurance activities discourage aggression in regions where potential security challenges exist. They include combat exercises, access agreements, competitor surveillance, and logistics. Many of these activities fall under the rubric of pre-positioning, in that they place the United States in a better position to win wars if deterrence fails. Access agreements and logistics/sustainment activities contribute to deterrence because they enable the United States to project power in regions far from U.S. territory, over an extended period of time. Combat exercises constitute a show of force to potential adversaries, build Army experience operating in the potential combat theater, and affirm solidarity with regional partner nations. Surveillance activities and partner intelligence cooperation improve the capacity of the Army to pre-position forces, anticipate competitor moves, and defuse potential security challenges before they gather momentum. Deterrence and reassurance can be viewed as a “low” relationship intensity strategy for shaping the theater. While it involves intense cooperation with the partner nations surrounding a target country, it requires almost no relationship-building with the focal country itself. When the U.S. relationship with a potential security challenger is poor, deterrence may be the first option for preventing the situation from deteriorating into war, and winning that war if deterrence fails.

Outreach activities plant the seeds of future partnerships with countries that currently have tenuous relationships with the United States. They include humanitarian assistance,

⁶ See, for example, bilateral US-Indonesia exercise Garuda Shield (“Garuda Shield: Exercise Garuda Shield,” n.d.); multi-lateral, AFRICOM-sponsored exercise Flintlock (United States Africa Command, “Flintlock,” n.d.).

⁷ O’Mahony et al., 2017b.

Figure A.1
Six Categories of Security Cooperation Activities



SOURCE: RAND Analysis of G-TSCMIS.

RAND RR2165A-A.1

activities to counter transnational threats, and international law enforcement. Humanitarian assistance provides a nonthreatening way for the Army to ameliorate basic sources of insecurity in a focal country, demonstrate the value of good relations with the United States, and catalyze informal channels for dialog between nations. Cooperation on transnational threats and law enforcement issues, such as smuggling, piracy, human trafficking, terrorism, and drug cartels, underscores that the United States and the focal country have common security interests, and that they can be better addressed working together than working alone. Outreach can be viewed as the “high” relationship intensity end of the theater-shaping spectrum. Ideally, the ultimate goal is to turn a security competitor or fragile state into a U.S. security partner, as well as a stabilizing influence on regional security.

Professionalization activities catalyze a military force that is capable of maintaining local order and professional discipline. They include basic military education with an emphasis on respecting human rights and maintaining military discipline. Each activity improves the ability of the partner military to nurture domestic security. Basic military education and training improve the partner military’s ability to be a force for domestic security, while stability operations diminish the destabilizing influences of entities that compete with the partner military for control. Emphasis on respecting human rights and on military discipline improves the chances that the partner military will be a positive element, rather than a liability in terms of coups, war crimes, and corruption. Professionalization can be viewed as the “low” relationship

intensity end of the capacity-building spectrum because it pragmatically strives to ensure that the local military does more good than harm for the local security situation.

Operational effectiveness activities improve the capacity of well-professionalized militaries and enhance their ability to be a stabilizing influence across their region. They include advanced military education, sales of sophisticated military equipment, and financial aid for partner militaries. All of these activities improve the ability of partners to project power across their region. Operational effectiveness can be viewed as the “high” relationship intensity end of the capacity-building spectrum because it empowers the partner military—trusting that the empowered partner will be a competent steward of American knowledge/technology and will behave in a manner compatible with American values.

Shared values and priorities activities build consensus on goals, methods, and plans, improving the likelihood that American and partner military actions will be politically compatible. They include senior leader visits, discussions on doctrine, and engagement with civilian political leadership. Senior leader visits demonstrate American commitment to partner nations and, along with political leader engagement, open space for reaching goal consensus. Discussions on doctrine, plans, and procedures translate mutual goals into practical implementation. Shared values and priorities can be viewed as the “low” relationship intensity end of the deepening partnership spectrum, because it starts from aspirations to improve a security partnership and strives to work out the practical aspects of how that partnership will work.

Joint operations activities develop capability to work closely together on operations. They include interoperability/standardization improvement, personnel exchanges, information-sharing, and joint research and development. Joint operations can be viewed as the “high” end of the deepening partnership spectrum, because each activity leverages the high degree of trust between partners to accomplish mutual objectives. Information-sharing and research/development involve entrusting partner nations with high-value information that could potentially damage national security if improperly used. Interoperability/standardization activities iron out the practical details of working together as a blended force, which potentially means depending on partner nation units under dangerous circumstances. Personnel exchanges go a step further, actively relying on partner military officers to contribute to the operations of U.S. Army units (and vice versa). The rewards for this trust are more knowledge of potential security threats, and more combined military resources with which to resolve them.

While each spectrum of security cooperation engagement is a unique dimension of engagement, different types of countries tend to fall at different intersections of the activity spectrums, depending on their military capabilities and the compatibility of their political values with American values. Moving clockwise along the figure, there is a natural progression from countries that have very low partnership depth to countries that are very close partners of the United States. In Figure A.1, the most common points of intersection are highlighted with gray lines.

When U.S. relations with a country are relatively weak, activity focuses on shaping the theater to reduce security risks. At first, this may simply involve deterrence and reassurance activities to contain potential threats, but then outreach activities can be used to plant the seeds of a U.S.-partner nation relationship. In moving from the “low” end to the “high” end of the theater-shaping spectrum, a country transitions from being a potential security risk to being a nascent partner. At this point, the United States can begin to improve the partner’s capacity to be a stabilizing influence. At first, this may simply involve enhancing the partner nation’s ability to maintain security in its own territory, and ensuring that the partner mili-

tary remains a positive influence on that security (“professionalization”). However, once the partner military has a demonstrated track record of being a stabilizing influence on domestic security (including demonstrated respect for human rights), it can be further empowered through operational effectiveness activities. In moving from the “low” to “high” end of the military capacity spectrum, a country transitions from being a potentially unreliable element in its domestic context to being a potential stabilizing element for the region. This paves the way for a country to become a close security partner of the United States. At first, this involves activities that build value consensus (“shared values and priorities”), ensuring that American and partner nation goals and methods align.⁸ Once strong consensus is achieved, the U.S. and partner nation militaries can begin operating jointly in ways that require trusting the partner nation to behave in ways that are not damaging to U.S. security. In moving from the “low” to “high” end of the deepening partnership spectrum, a country transitions from being willing to find common ground to conducting joint operations that depend on that common ground.

The gray line across the center depicts a rare “shortcut” path, compared with the more typical clockwise progression. Some countries already have strong military capability but low political compatibility. Such countries may have security goals that are highly incompatible with American values and/or compete with the United States to shape the security environment. For such countries, there is no need to traverse the capacity-building segments of the scale. Instead, Army activities focus on demonstrating compatibility of interests through *outreach* activities and building compatibility of values through *shared values and priorities* activities.

⁸ Shared values and priorities activities are also useful when a partner nation has the military capacity of a country eligible for “effectiveness” activities but the political cooperativeness of a country suited for “outreach” activities. This is depicted with a dashed gray line in Figure A.1.

Regression Model Methodology

This appendix is written for a reader with a basic understanding of regression analysis. It provides more detail into the linear modeling approach used to obtain the empirical results in Chapter Three. Each table reports how much more/less likely countries with given characteristics are to attend security cooperation activities. The numbers are odds ratios, so numbers greater than one indicate that a country with a given characteristic is *more* likely than the average country to attend security cooperation activities, and numbers less than one indicate that a country is *less* likely than the average country to attend security cooperation activities. For example, the coefficient for NORTHCOM in “control” column of Table B.1 is 7.552. That means that if the average country had a 1 percent chance of attending each event, then CENTCOM countries would have a $1 \times 7.552 = 7.552$ percent chance of attendance.

Chances of Attending Security Cooperation Activities

Table B.1 presents results for activity participation. To obtain the attendance findings, we treat each country’s involvement (or noninvolvement) at each security cooperation activity in G-TSCMIS as a separate observation, yielding 2.63 million observations between 2006 and 2014. We then examine which country characteristics most accurately predicted country involvement. The column highlighted in blue (“Best Fit”) indicates the final model used to calculate the statistics reported in the chapter. The other columns report the models used to confirm that Best Fit was accurate (low BIC score, indicating good predictive power), parsimonious (uses as few coefficients as needed), and robust (each row of coefficients consistently indicates either an increase or a decrease in chances, and is statistically significant).

Control—The model on the far left measures how much activity involvement can be predicted with factors that have no relationships to the countries themselves, variation in activity tempo over time, and differences in GCC engagement patterns. If country characteristics do not really correspond to differences in engagement, then subsequent models will fail to explain the data any better than this one. In essence, this model determines the minimum that a model must pass to be a good model of activity involvement. The BIC row at the bottom quantifies that minimum threshold. Lower BIC scores indicate a better model.

All Factors—The model on the far right measures how much activity involvement can be predicted using all of the country characteristics we collected. If the characteristics dataset has been chosen well, this model will have the best BIC score, but it will also be the least

Table B.1
Activity Participation

		Control	Model 1	Model 2	Model 3	Model 4a	Model 4b	Best Fit	All Factors
Base Rates	Average rate (intercept)	0.008***	0.001	0.005***	0.007***	0.002***	0.006***	0.007***	0.000***
	Multilateral event		1.115***	1.115***	1.115***	1.115***	1.115***	1.115***	1.116***
Data Quality Controls	Year2007	1.083**	1.051^	1.056*	1.061*	1.065*	1.061*	1.058*	1.056*
	Year2008	1.609***	0.944*	0.947*	0.948*	0.943*	0.943*	0.944*	0.942*
	Year2009	1.278***	0.967	0.972	0.975	0.977	0.974	0.97	0.965
	Year2010	1.191***	1.054*	1.050*	1.048*	1.038^	1.041^	1.038^	1.032
	Year2011	1.351***	1.128***	1.124***	1.107***	1.1***	1.098***	1.095***	1.093***
	Year2012	1.287***	1.052*	1.048*	1.037^	1.028	1.02	1.026	1.018
	Year2013	1.242***	1.098***	1.090***	1.074**	1.072**	1.048*	1.066**	1.058*
	Year2014	1.200***	1.078**	1.078**	1.056*	1.086*	1.04	1.043^	1.038
GCC	CENTCOM	1.399***	1.117***	1.260***	1.307***	1.542***	1.410***	1.436***	1.504***
	EUCOM	1.788***	1.053*	1.036	0.861***	1.108	0.933**	0.880***	0.861***
	NORTHCOM	7.552***	5.248***	4.707***	3.897***	5.553***	3.818***	3.868***	3.798***
	PACOM	3.551***	2.394***	2.412***	2.158***	3.016***	2.209***	2.200***	2.103***
	SOUTHCOM	1.836***	1.542***	1.586***	1.298***	1.418***	1.324***	1.314***	1.233***
Strategic	Econ. power (GDP)		1.034**	1.011	0.970***				1.014
	Supportiveness in UN					0.998*	0.995***	0.987***	1.634***
	Wealth (GDP per capita)		0.951***	0.868***	0.947***	1.083***	0.955***	0.962**	0.963**
Political	Human rights respect		1.491***	1.278***					1.156***
	Democratic (Polity IV)				1.027***	1.027***	1.028***	1.026***	1.024***
	Recent coups		1.155***						1.176***
	State fragility index (SFI)			0.381***	0.376***	0.893*	0.298***	0.345***	0.272***
Cultural	English speakers		1.217***	1.429***	1.323***	0.934***	1.340***	1.314***	1.245***
Organizational	Military technical skill								0.774***
	Military professionalism		1.002***	0.998**	0.998**	1.007**	0.998**		1.002**
	Population education		1.939***	1.670***	1.590***	1.371***		1.920***	1.942***
	Military resources		1.018***	1.02***	1.022***		1.019***	1.020***	1.021***
Aid	Amount of assistance		1.001	1.002^	1	1.005***	1	1	1.001
	Assistance inconsistency		1.133***	1.134***	1.125***	1.077***	1.123***	1.122***	1.11***
	Bayesian information criteria (BIC)	494,534	434,132	434,089	433,713	435,988	433,848	433,575	433,091
	N	2.63m	2.63m	2.63m	2.63m	2.63m	2.63m	2.63m	2.63m

SOURCE: RAND analysis of G-TSCMIS data on security cooperation activities.

* P < .05; ** P < .01; *** P < .001

efficient model. In essence, this model determines the maximum BIC improvement that country characteristics might provide, relative to the Control model, but tells us the least about which characteristics are truly important for understanding engagement patterns.

Models 1–4b—The middle models, excluding Best Fit, explore how well different combinations of characteristics predict activity involvement. Moving from left to right, each model builds on the insights of the last,¹ producing progressively better models and double-checking each other's assessment of the importance of characteristics. These model progressions are especially important when examining country data because such data are notorious for producing inconsistent results across model specification.² For a model result to be trustworthy, it should have, at a minimum, consistent sign and statistical significance across most model specifications.

Best Fit—The highlighted (in blue) model reports a model that, given a preponderance of the other models, is accurate, parsimonious, and robust. The model answers the research question, “What kinds of countries are more likely to have been involved in security cooperation activities?”

Portfolio Share of Six (Nonexclusive) Activity Types

Table B.2 presents results for portfolio share by activity type. To obtain the portfolio share models, we use the same model specifications as attendance Best Fit, but instead apply them toward predicting activity characteristics. For these models, there are 52,689 observations, one for each instance of a country attending a G-TSCMIS activity. The six models predict whether those activities fell into each of six (nonexclusive) activity categories, given the characteristics of the involved country. Each model answers the research question, “What kinds of countries tend to have a larger proportion of their engagement portfolio devoted to activities of this type?”

¹ Model progressions: many of the variables in these models correlate strongly with each other. For example, the number of “recent coups” in a country correlates strongly with the “state fragility” of that country. Each model progression substitutes a variable for one closely related to determine which produces a model that more accurately predicts the chances that a country will attend security cooperation activities. For example, model 1 uses “human rights respect” and “recent coups” to measure how political characteristics correspond to differences in event attendance. However, by model 3, “democratic” and “state fragility” are used instead, because they more accurately predict activity participation. Likewise, strategic characteristics are measured with “econ. power (GDP)” and “income (GDP per capita)” in model 1, but by model 4a they are measured with “supportiveness in UN” and “income (GDP per capita).”

² Cristobal Young, “Model Uncertainty in Sociological Research: An Application to Religion and Economic Growth,” *American Sociological Review*, Vol. 74, No. 3, 2009.

Also see these examples of the pitfalls of model instability:

William Easterly, Ross Levine, and David Roodman, “New Data, New Doubts: A Comment on Burnside and Dollar’s ‘Aid, Policies, and Growth’ (2000),” National Bureau of Economic Research Working Paper No. 9846, July 2003; Craig Burnside and David Dollar, “Aid, Policies, and Growth: Reply,” *The American Economic Review*, Vol. 94, No. 3, 2004.

Table B.2
Portfolio Share by Activity Type

		Outreach	Professionalization	Effectiveness	Values and Priorities	Joint Operations	Deterrence and Reassurance
Base Rates	Average rate (intercept)	0.097***	3.270***	0.000***	0.109***	0.115***	0.022***
	Attendance chances	1.027***	1.068***	0.623***	1.074***	0.923***	0.892***
Data Quality Controls	Year2007	1.369***	1.104	0.551**	1.002	1.135^	0.710***
	Year2008	0.923	2.123***	2.028***	1.007	0.814**	0.492***
	Year2009	0.918	1.513***	3.743***	0.906^	0.834**	0.438***
	Year2010	1.595***	1.02	3.323***	0.955	1.09	0.485***
	Year2011	1.579***	0.815***	2.225***	0.867*	1.650***	0.496***
	Year2012	1.108	0.953	2.988***	0.693***	1.126^	0.470***
	Year2013	0.954	1.065	2.507***	0.974	1.515***	0.523***
	Year2014	1.495***	1.406***	2.440***	1.025	0.984	0.614***
GCC	CENTCOM	1.318***	0.766***	2.639***	0.819**	3.539***	0.707***
	EUCOM	1.048	0.622***	0.983	0.901	2.817***	0.311***
	NORTHCOM	1.983***	0.505***	1.843***	0.602***	3.454***	0.541***
	PACOM	4.354***	0.478***	0.817^	0.876*	1.940***	0.895^
	SOUTHCOM	0.620***	0.670***	0.838	0.947	1.685***	0.251***
Strategic	Econ. power (GDP)						
	Supportiveness in UN	0.988***	1.001	0.996	0.996	1.019***	0.993^
	Wealth (GDP per capita)	1.121**	0.863***	1.261***	1.129***	0.801***	1.230***
Political	Human rights respect						
	Democratic (polity IV)	0.987***	1.004	1.048***	1.002	1.013**	1.007^
	Recent coups						
	SFI	5.004***	0.740*	2.500***	1.132	0.198***	0.964
Cultural	English speakers	0.862**	1.011	1.009	0.820***	1.490***	0.880*
Organizational	Military technical skill						
	Military professionalism						
	Population education	0.666***	1.273**	0.961	0.766**	0.889	2.730***
	Military resources	0.990***	0.996***	1.008***	0.999	1.002*	1.009***
Aid	Amount of assistance	0.991***	1	1.001	1.009***	0.999	0.998
	Assistance inconsistency	0.995	0.912***	1.063^	1.061***	1.016	1.007
BIC		35,837	54,853	15,748	36,853	40,990	32,784
N		52,689	52,689	52,689	52,689	52,689	52,689

SOURCE: RAND analysis of G-TSCMIS data on security cooperation activities.

* P < .05; ** P < .01; *** P < .001

Dimensions of Partner Capabilities Assessment

This appendix discusses the six dimensions of analysis to include in a partner capabilities assessment. These dimensions are

- security environment
- objectives
- capabilities
- characteristics
- will
- previous engagements, particularly security cooperation activities.

Partner Security Environment

Understanding a partner's security environment is crucial for understanding the context of all the previous components. It affects what the partner's objectives are, how leveraged its capabilities are, and how committed it is to working with the United States to achieve shared security cooperation objectives. A partner's security environment will provide Army planners with an understanding of what types of security cooperation are most valuable to the partner and identify constraints on what might be effective.

As with understanding partners' objectives, assessing partner countries' security environments is a key task for the Senior Defense Official/Defense Attache. Open-source information is also available through news sources. A good database on low-level conflicts that may escalate is the Heidelberg Institute for International Conflict Research's Conflict Barometer.¹

Partner Objectives

A critical component of a partner assessment is identifying what the partner's objectives are in participating in a security cooperation activity. Just as the ASCC develops priority areas based on perceived opportunities for growth or in response to a crisis or concern, partners bring their priorities, goals, and concerns to the table and should be consulted as much as possible. Not only does assessing partner objectives ensure that the Army is properly matching activities to

¹ Heidelberg Institute for International Conflict Research, *Conflict Barometer*, n.d.

partner priorities, it also provides partner nations with more ownership over the activity, which is one of the key indicators of whether an activity will be absorbed and sustained long after the Army is actively involved in an implementing role.²

At a minimum, the Army should look to U.S. government reports on partner nation objectives, such as those that come from the DoS on diplomatic cooperation and annual agreements or official statements from the Security Cooperation Organization. The Senior Defense Officials/Defense Attache stationed at U.S. embassies abroad are also excellent sources of information on partner priorities and goals, as they interact with representatives from partner nations on a regular basis and have their fingers on the pulse of their national and military objectives and strategies as well as GCC TCP objectives. Where possible and appropriate, direct coordination with a partner on objectives and security cooperation priorities provides a more comprehensive perspective of partner objectives. The Army already achieves this kind of deeper coordination with some partners, for example, ARNORTH activity with Mexico.

Partner Capabilities

Accurately assessing a partner's capability is essential to measuring absorptive capacity, and as numerous studies discussed in Part I of this report have pointed out, properly matching absorptive capacity to security cooperation activities is both critical to activity success and currently poorly managed. Assessing partner capability means examining partner capabilities across the DOTMLPF-P domains. Much of the assessment information necessary to minimally assess partner capabilities is available in open-source materials, such as those provided by Jane's Defense and Security Intelligence and Analysis³ and in U.S. government reports. RAND has also developed several assessment tools to examine functional areas across DOTMLPF-P domains.

Partner Characteristics

In addition to understanding the specific military context of a partner, ASCC planners also need to understand a partner's broader context, including relevant societal norms, demographic trends, economic and political interests, and influential actors, among others. Understanding partner characteristics matters because they affect how security cooperation LOEs are implemented, whether they receive public support or pushback, how well the partner nation will be able to sustain the effort with reduced or no U.S. assistance after a given time period, and whether the United States will ultimately be making a worthwhile investment in the partner's long-term stability and security.

Creating a systems profile of partner characteristics can be done with broad strokes, when only a minimal assessment is necessary, or it may involve qualitative and quantitative data from SME analysis based on structured observation and interaction. The PMESII system was developed to provide a systemic picture of a partner country and its various components.

² Paul et al., 2013a.

³ Jane's offers a range of defense analysis products, as described at www.janes.com.

Multiple organizations track indicators on elements of the PMESII framework, and those indicators are open source and updated regularly, including the World Bank, the UN, Stockholm International Peace Research Institute, and many others. At the higher end, organizations such as the Institute for Economics and Peace produce yearly indices of country stability and resilience toward internal and external shocks, which combine multiple indicators into a single weighted score.⁴

Partner Will

Gauging the willingness of a partner to participate in security cooperation activities and to follow up that participation with institutional reforms or other forms of progressive measures, and, crucially, gauging the willingness of its security forces to fight and die in defense of their objectives is a critical component of a partner assessment. *Willingness* may take many different forms, from putting forth matching financial contributions for an activity, to releasing statements of support, to signing access agreements with the United States. Some partners may, based on their political preferences or historical relationship with the United States, be more naturally inclined to participate in security cooperation activities with the Army. Others may qualify as nascent security cooperation partners and use security cooperation activities to signal their willingness to deepen and solidify the current relationship. Willingness to participate may shift dramatically in either direction over time, making it all the more important to have some way to assess and track patterns to make informed decisions about when and how to collaborate.

While reliable information on willingness can be difficult to track down, the assessments created by GCCs and country teams offer some basic level of observation on partner will. At a more comprehensive level, ASCCs can produce objective-specific analyses of partners, which include assessment of partner will for a particular activity or LOE.

Previous Engagements

One of the best indications of future security cooperation engagements is a history of past engagement. Tracking previous security engagements the Army has undertaken is important for a number of reasons: first, it encourages transparency and accountability around the personnel, equipment and supplies, and budget used to support previous efforts. Second, tracking previous engagements shows the trajectory and pattern of a security cooperation relationship: if the Army conducts the same kinds of activities year after year with one country, with little or no measured improvement in the desired outcome, this should be a strong indicator that either partner will or capabilities is insufficient to create desired effects.

The ability to track engagement history has improved significantly with the advent of the G-TSCIMIS system, and will continue to evolve and improve as more credible data are added

⁴ See, for example, the Global Peace Index, published by the Institute for Economics and Peace, which includes sub-components analyzing internal and external conflict, societal safety and security, and a country's militarization. Institute for Economics & Peace, *Global Peace Index, 2016: Ten Years of Measuring Peace*, Institute for Economics & Peace IEP Report 30, June 2016.

to the system. Outside of G-TSCIMIS, AARs can provide useful insights into successes and failures of project implementation and possible opportunities for future engagements. At a deeper level, program implementers should consider creating data repositories for activity-level information on AM&E that is linked back to ASCC tasks to track longitudinal progress of those activities and make midcourse corrections when necessary.⁵

⁵ See Watts et al., 2017, pp. 33–58 for a more in-depth discussion of how to develop data repositories relevant to security cooperation.

Examples of Three Notional Activity-Level TOCs

This appendix presents notional examples of three security cooperation activities. The information included in them is simplified, but the differences across the three examples highlight how the difficulty of creating informative TOCs reflects different activity characteristics.

- Easy to establish TOC: Army Civil Affairs provides MEDCAP to Djibouti village
- Medium difficulty TOC: Colombian participants in MCCC at Fort Benning (WHINSEC)
- Hard to establish TOC: Southern Accord 2016 (Malawi).

MEDCAP to Djibouti Village

- IMO: Improve rural Djibouti's health care system
- Difficulty: easy to establish.

Initial Assessment for Activity Planning

Capabilities and Characteristics

- Djibouti's health system faces severe challenges, including a lack of human capital in health services and a lack of funds for health programs.
- The Djiboutian government created the Health Development Plan 2013–2017, which is intended to enhance partnerships and develop institutional capacity.

Potential Risks and Indirect Effects

- Medical personnel providing MEDCAP could get infected with new or severe local diseases that are difficult to treat.
- Improving the health status of patients who receive the treatment directly could have positive spillovers for the Djiboutian population who do not receive the treatment directly. People in Djibouti become less sick in general.

Contextual Data

- Djibouti requires more trained physicians—currently there are 0.229 physicians for every 1,000 people.
- The goal is to provide health care services to as many patients as possible within the two days of MEDCAP, and to ensure Djiboutian physicians get trained on best practices for preventive care.

Relevant PMESII Information

- Djibouti has a number of remote, dispersed communities with unreliable transportation and poorly maintained roads. Forty percent of the population is below the poverty line, and Djibouti has not made significant progress on Millennium Development Goals, including the health-related ones.

Other Stakeholders

- Djibouti receives health assistance from numerous international organizations and NGOs including the UN, World Bank, USAID, and multiple others, and civil engineering teams from the Army are working with teams from Sudan, China, Italy, and others to improve road and health infrastructure.

TOC

Below, we present the components for this TOC in list form. They are shown graphically in Figure D.1.

Inputs

- Physicians from Combined Joint Task Force-Horn of Africa (CJTFHA) and USAID
- Djiboutian physicians
- Cuban physicians
- Patients
- Pharmaceutical and medical supplies
- Medical equipment (e.g., heart monitors, stethoscope)
- Clinical facility.

Why will inputs support activities?

- Physicians will provide medical support and treatment.
- Pharmaceutical and medical supplies are required for physicians to provide support and treatment.

Activities

- Physicians from CJTFHA, Djibouti Ministry of Health, USAID, and Cuba provide medical care in Yoboki, Djibouti, and training to Djiboutian physicians through comparing diagnoses for illnesses.

Why will activities support outputs?

- American, Cuban, and Djiboutian physicians providing medical care to patients and comparing diagnoses for illnesses supports Djiboutian individuals receiving treatment and Djiboutian physicians receiving training.

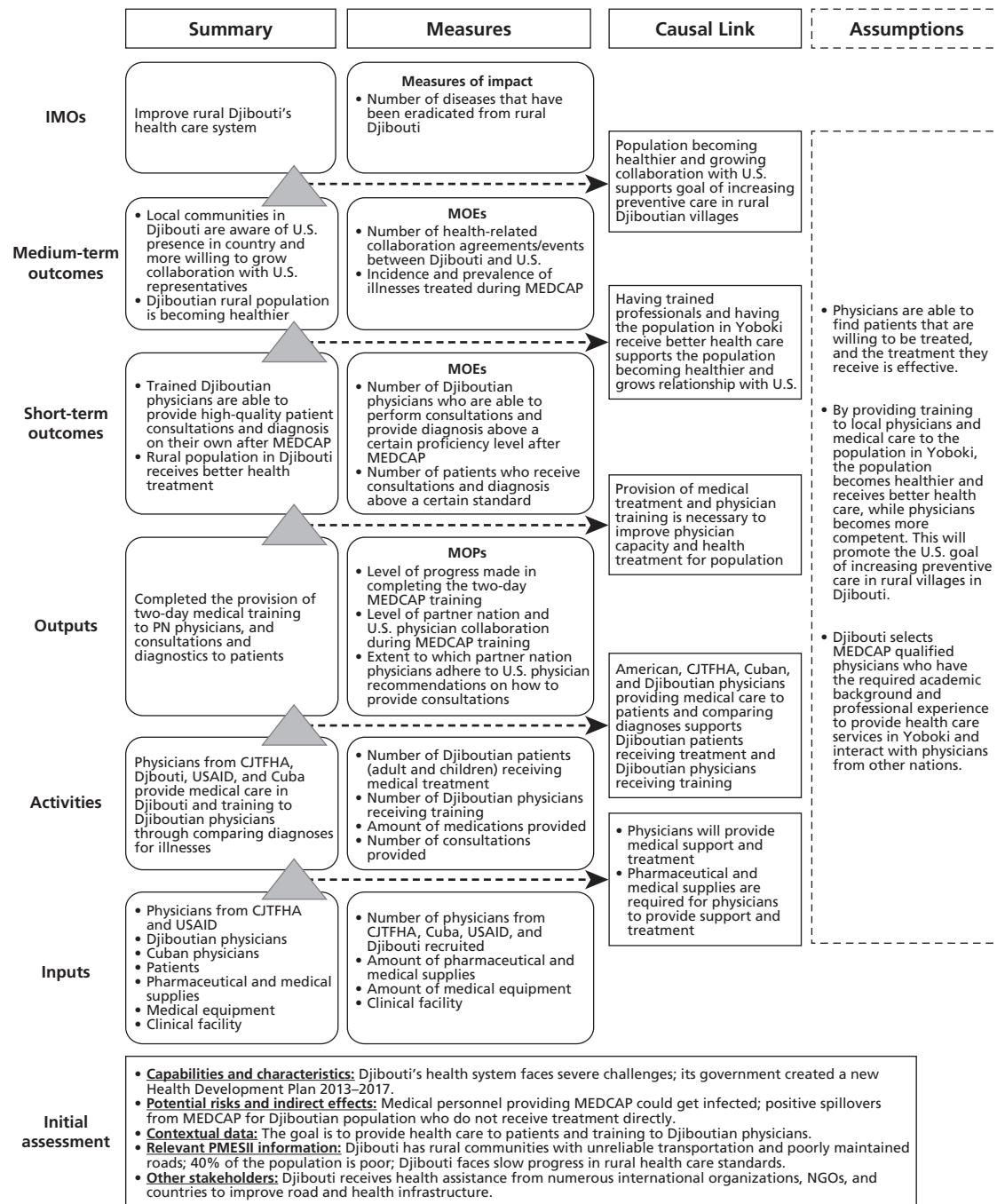
Outputs

- Patients, including children, received medical treatment and medications for their illnesses.
- Djiboutian physicians completed two-day MEDCAP training.

Why will outputs support outcomes?

- Provision of medical treatment and physician training is necessary to improve population's health treatment and physician capacity.

Figure D.1
Notional TOC for MEDCAP to Djibouti Village



SOURCE: RAND analysis.

RAND RR2165A-D.1

Outcomes¹

- Trained Djiboutian physicians are able to perform consultations and provide diagnoses on their own after MEDCAP.
- Rural population in Djibouti receives better health treatment, including preventive health care.
- Local communities in Djibouti are aware of U.S. presence in country and more willing to grow collaboration with U.S. representatives.
- Djiboutian rural population is becoming healthier.

Why will outcomes support IMO?

- By having trained professionals, the population in Yoboki will receive better health care and become healthier, which will grow U.S.-Djibouti collaboration support goal of increasing preventive care in rural Djiboutian villages.

IMO

- Improve rural Djibouti's health care system.

Activity Assumptions

- Physicians are able to find patients that are willing to be treated, and the treatment they receive is effective.
- By providing training to local physicians and medical care to the population in Yoboki, the population becomes healthier and receives better health care, while physicians become more competent. This will promote the U.S. goal of increasing preventive care in rural villages in Djibouti.
- Djibouti selects MEDCAP qualified physicians who have the required academic background and professional experience to provide health care services in Yoboki and interact with physicians from other nations.

Colombian Participants in MCCC

- IMO: Colombian combined-arms staff officers and commands are capable of performing combat tasks in a full-spectrum environment against an adaptive enemy
- Difficulty: medium to establish.

Initial Assessment for Activity Planning***Capabilities and Characteristics***

- U.S. security cooperation activities with Colombia have focused largely on drug-related operations, but as Colombia's military becomes more capable and professionalized, its military personnel will require support for professionalization and expertise in other areas.
- Colombia has been a longtime U.S. security cooperation partner and has participated in WHINSEC courses at Fort Benning and other locations for many years.

¹ We do not differentiate between short- and medium-term outcomes in the main text, but we do differentiate between them in Figure D.1. Practitioners may use both approaches

Potential Risks and Indirect Effects

- U.S. military personnel may not be able to track how trained army personnel use the MCCC content after completion; hence, potential threats originate from trained army personnel using the knowledge to conduct activities unrelated to MCCC goals in Colombia.
- MCCC might increase United States' vulnerability to foreign attacks as foreign military personnel increase their knowledge of U.S. operations.
- Training Colombian staff officers can have positive spillovers as trained officers may share their knowledge with other Colombian military personnel who did not participate in MCCC.

Contextual Data

- Previous tests conducted with Colombian participants established that 50 percent lacked training to adequately perform combat tasks.
- At the end of the course, students are expected to be more creative, adaptive, agile, and self-confident combined-arms staff officers. They are also expected to be capable of performing combat tasks in a full-spectrum environment against an adaptive enemy.²

Relevant PMESII Information

- The environment or conditions in which the training takes place are especially designed for this type of training. Training takes place at the U.S. Army's Maneuver Center of Excellence in Fort Benning, Georgia, and students stay at student facilities or at hotels in Fort Benning.

Other Stakeholders

- Military personnel from other partner nations (other than Colombia) also participate in the training, as MCCC is delivered to military personnel from different partner nations simultaneously.

TOC

Below, we present the components for this TOC in list form. They are shown graphically in Figure D.2.

Inputs

- Colombian Army officers participating in MCCC
- U.S. team chief
- U.S. Army small group leaders
- Training materials (e.g., notebooks, computers)
- Transportation and accommodations for a 3-day course in Washington, D.C., and a 22-week course at Fort Benning, Georgia.

Why will inputs support this activity?

- Students, instructors, and training material provided with the given time period are all necessary for the course to occur.

² Western Hemisphere Institute for Security Cooperation, 2015, p. 11.

Activity

- MCCC–WHINSEC is delivered in Fort Benning, Georgia, and Washington, D.C.

Why will this activity support outputs?

- The course is necessary for students to receive certification and will contribute to professionalization of the foreign military participants.

Outputs

- Officers from Colombia successfully completed and passed a 22-week course, which required
 - creating company-level organizational plans for all staff-level personnel for future participation in full-spectrum activities
 - completing one class simulation (topic may vary)
 - completing training on human rights, rule of law, rules of engagement, due process, civilian control of the military, and the role of the military in society
 - completing three days of discussions in briefings with representatives from federal agencies and NGOs in Washington, D.C.

Why will outputs support outcomes?

- Colombian officers participating in the MCCC will lead to increased confidence and professionalization to perform combat tasks in a full-spectrum environment against adaptive enemies.
- Simulation exercise and training interactions between Colombian students and U.S. Army staff will lead to enhanced understanding of operational culture and doctrine.
- Trip to Washington will increase professional connections with U.S. agencies and other regional organizations and promote relationship-building.

Outcomes

- Increased capability of Colombian officers while performing combat tasks in a full-spectrum environment against adaptive enemies
- Enhanced understanding of operational culture and doctrine for Colombian officers and U.S. Army
- Increased professionalization and career confidence of Colombian combined-arms staff officers who attend the MCCC
- Enhanced relationships between Colombian officers and American counterparts through interactions.

Why will outcomes support IMOs?

- Increased professionalization of staff-level officers will help support the overall capability of the Colombian army.

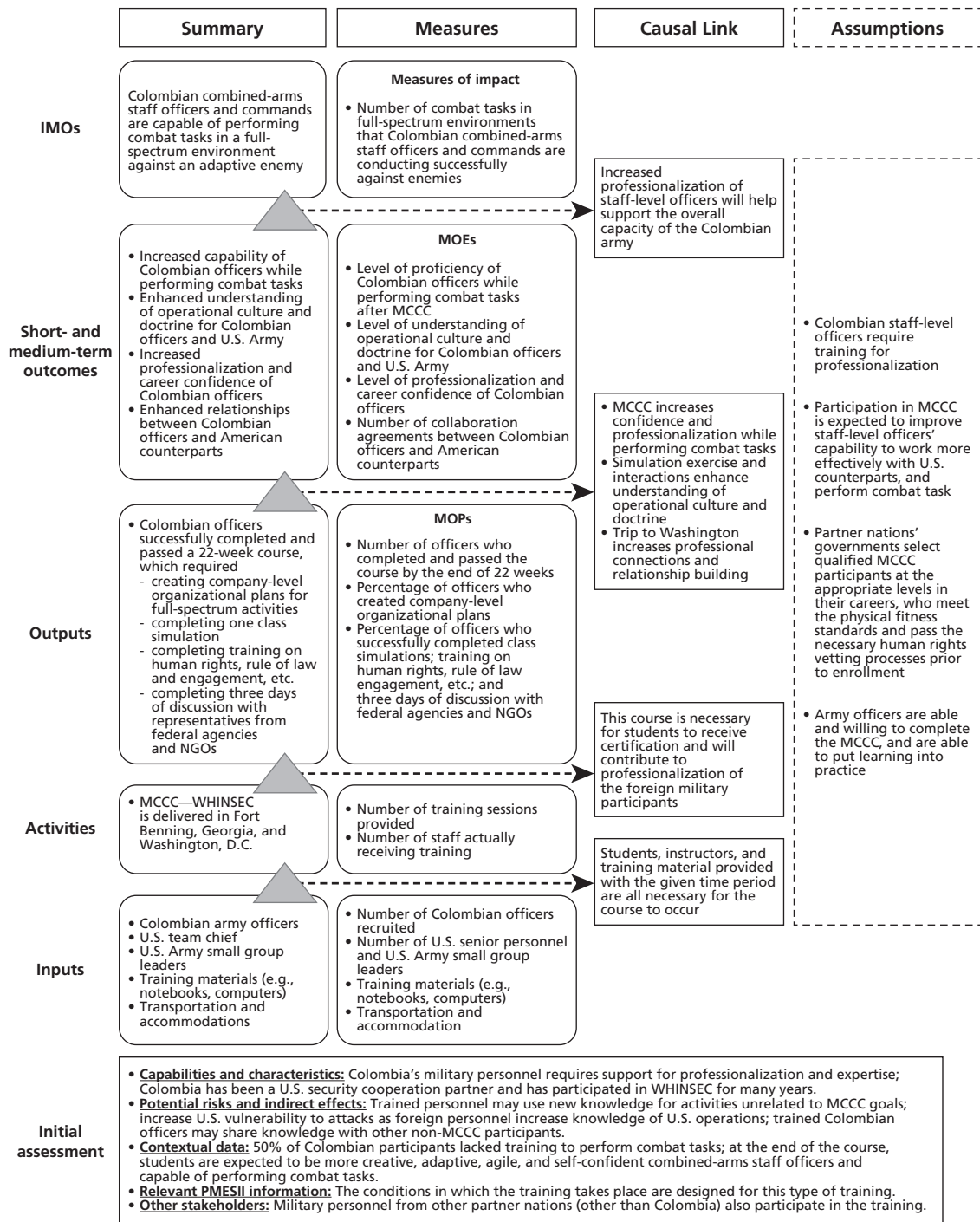
IMO

- Colombian combined-arms staff officers and commands are capable of performing combat tasks in a full-spectrum environment against an adaptive enemy.

Activity Assumptions

- Colombian staff-level officers require training for professionalization.
- Participation in MCCC is expected to improve staff-level officers' capability to work more effectively with U.S. counterparts and perform combat tasks.

Figure D.2
Notional TOC for Colombian Participants in MCCC



SOURCE: RAND analysis.

RAND RR2165A-D.2

- Partner nations' governments select qualified MCCC participants at the appropriate levels in their careers who meet the physical fitness standards and pass the necessary human rights vetting processes prior to enrollment.
- Army officers are able and willing to complete the MCCC and are able to put learning into practice.

Southern Accord 2016 (Malawi)

- IMOs: (1) increased capacity of national military and civilian sectors in Malawi to respond to humanitarian and natural disasters; (2) enhanced cooperation within the Southern African Development Community (SADC)
- Difficulty: difficult to establish.

Initial Assessment for Activity Planning

Partner Capabilities and Characteristics

- Malawi has recognized the importance of establishing a unified policy to prepare for and respond to humanitarian and natural disasters, and did so in the recent national disaster.
- Malawi currently lacks a system for identifying, assessing, monitoring, and mapping disaster risk at all levels.

Potential Risks and Indirect Effects

- Training on systemwide coordination in disaster management could spill over into demand for greater government-wide coordination on other topics.
- The severity of the disaster may affect the safety of military personnel and jeopardize training effectiveness, as humanitarian aid activities tend to occur in places with other existing conflicts or facing chronic adversities.³

Information to Guide Initiative Design

- After participating in Southern Accord, Malawian military personnel and civilian officials will understand how to coordinate civil-military efforts on disaster response.

Contextual Data

- Before Southern Accord, only 20 percent of military personnel had adequate humanitarian and natural disaster response training.
- The goal of the ten-day exercise is to increase interoperability and knowledge of the SADC on disaster response.

Relevant PMESII Information

- Eighty percent of the Malawian population is rural and depends on subsistence-based farming for their livelihoods, which makes them very vulnerable to any droughts or floods.
- Natural disasters are seen by the Malawian government as a key factor in inhibiting economic growth and poverty reduction.

³ Colleen A. McFarlane, "Adjustment Risks of Humanitarian Aid Workers," *The Australasian Journal of Disaster and Trauma Studies*, Vol. 2004-1, 2004.

Other Stakeholders

- Malawian civil society groups, the United Nations, and nongovernmental humanitarian organizations are all involved in the effort to coordinate disaster and humanitarian response in Malawi.

Activity Assumptions

- Southern Accord is effective at increasing the capacity of military personnel to respond to humanitarian and natural disasters, and has the potential to improve cooperation among implementing partners.
- Malawian government selects personnel with a background in humanitarian/disaster relief operations, peacekeeping operations, and aeromedical evacuation to participate in Southern Accord.
- The government of Malawi and Malawian military personnel are receptive to the exercise, as they demand training and support to adequately respond to natural disasters.
- African forces are able and willing to complete the exercise, and can successfully put knowledge into practice.
- Differences between civilian and military roles in humanitarian and disaster response (i.e., who leads/who supports) between the United States and Malawi would not be consequential to the conduct of this tabletop exercise.

TOC

Below, we present the components for this TOC in list form. They are shown graphically in Figure D.3.

Inputs

- USARAF planners
- Military planners from Malawi
- Military members from SADC
- Members from the Seventy-Fifth Training Command
- Land Operations Support Command of the Royal Netherlands Army
- Supplies for tabletop exercise (e.g., emergency equipment)
- Civilian agency participants
- Infrastructure (e.g., classroom, conference rooms).

Why will inputs support activity?

- Personnel from military and multinational offices, as well as officials from SADC, using exercise supplies are necessary for successful implementation of Southern Accord.

Activity

- Command post and tabletop exercise takes place in Salima, Malawi.

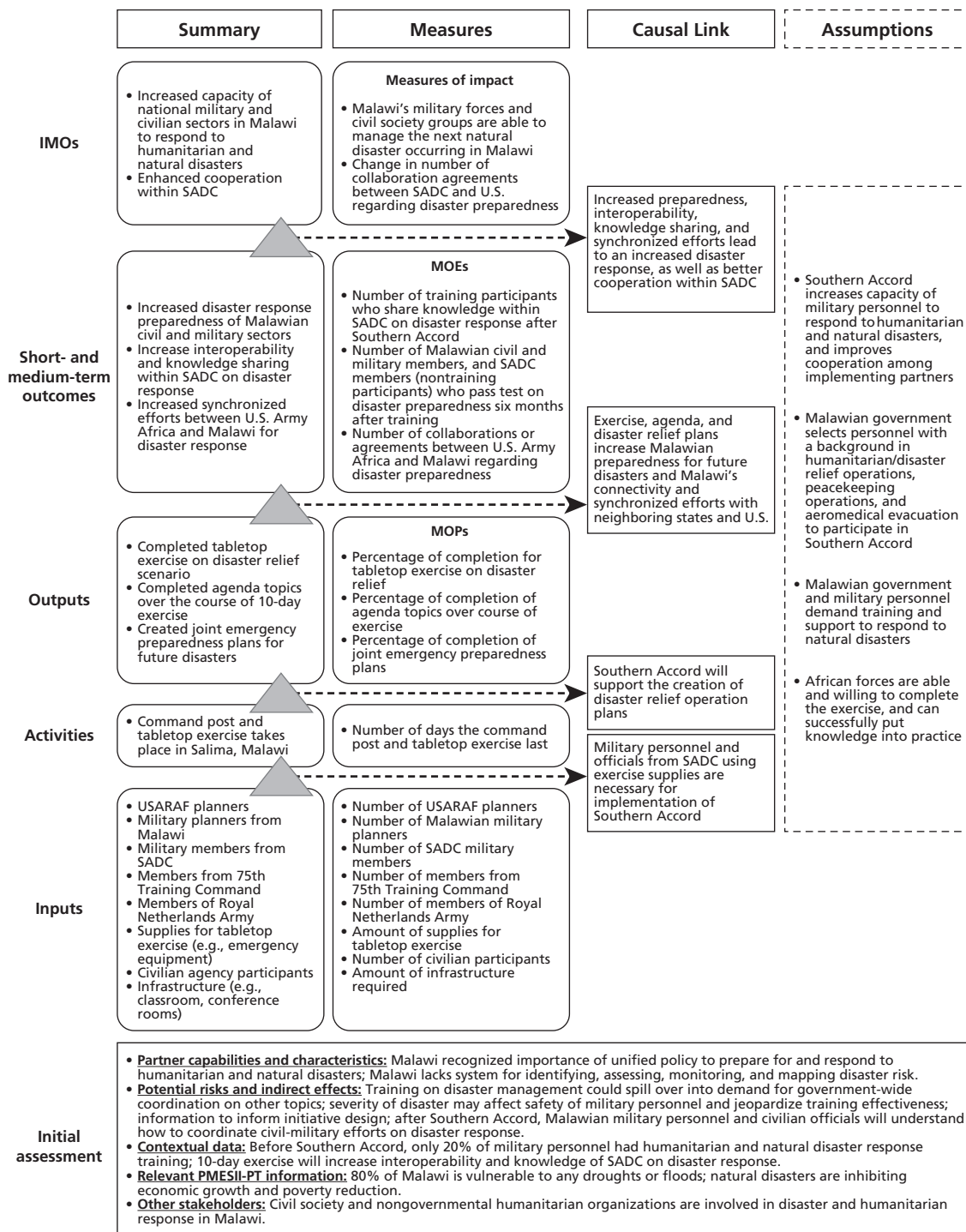
Why will this activity support outputs?

- Southern Accord will support the creation of disaster relief operation plans.

Outputs

- Completed tabletop exercise on disaster relief scenario
- Completed agenda topics over the course of ten-day exercise
- Created joint emergency preparedness plans for future disasters.

Figure D.3
Notional TOC for Southern Accord 2016 (Malawi)



SOURCE: RAND analysis.

RAND RR2165A-D.3

Why will outputs lead to outcomes?

- Completed exercise, agenda, and disaster relief plans will increase Malawian preparedness for future disaster scenarios and increase connectivity to neighboring states, as well as synchronized efforts between the United States and Malawi.

Outcomes

- Increased disaster response preparedness of Malawian civil and military sectors
- Increased interoperability and knowledge-sharing within SADC on disaster response
- Increased synchronized efforts between U.S. Army Africa and Malawi.

Why will outcomes lead to IMOs?

- Increased preparedness, interoperability, knowledge-sharing, and synchronized efforts will lead to an increased capacity for future disaster response, as well as better cooperation within SADC.

IMOs

- Increased capacity of national military and civilian sectors in Malawi to respond to humanitarian and natural disasters
- Enhanced cooperation within SADC.

Activity-Level AM&E Decision Tree and Supporting Annexes

This appendix elaborates on the decision tree presented in Chapter Six and presents supporting annexes to help determine appropriate AM&E approaches for security cooperation activities. Figure E.1 maps the paths of each of the three example activities used in Chapter Six and Appendix D through the decision tree. The decision tree is designed to guide security cooperation planners through a process for determining the type of AM&E an activity warrants, based on activity characteristics and what we learned from previous activities. These annexes supplement the template by providing brief explanations for concepts raised along that decision tree. The annexes are not designed to be comprehensive or definitive for each topic. Instead, they are meant to provide guidance to the planner about (1) how to work through the template and (2) where to seek additional or more detailed information.

Annex A: Activity-Level TOC

What Is an Activity-Level TOC?

An activity-level TOC illustrates the logic for how an activity is expected to accomplish the planner's desired objectives in a particular environment or setting in which the activity takes place.¹ The best time to design a TOC is at the planning stage for an activity, since at this stage planners identify long-term objectives and the potential activities to accomplish those objectives.

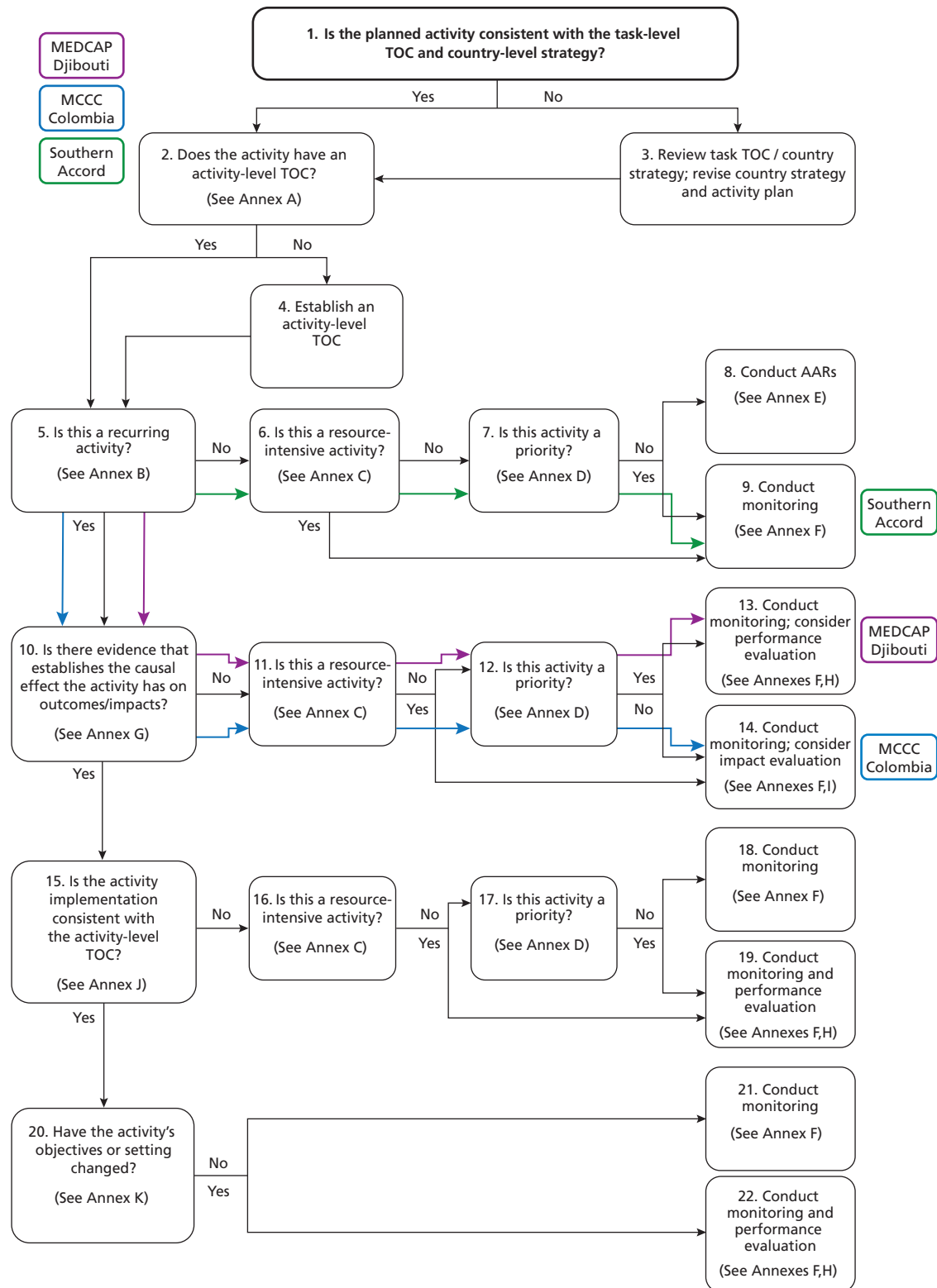
Why Is It Useful?

TOCs help map out activities, identify side effects from an activity, establish monitoring and evaluation needs, and provide the basis for measuring results. By providing a detailed roadmap of how change happens (i.e., what outcomes the planner wants to achieve and how to get there), TOCs enable activity planning, monitoring, and evaluation. A TOC also identifies how changes in outcomes and long-term objectives are expected to be achieved through the activity.

A single TOC is static: it captures how an activity is expected to work in a particular setting with a specific set of objectives. However, activities change over time as they are adapted to suit their specific environment. As a result, TOCs may need to be changed and updated over time to adjust to settings or objectives.

¹ Center for Theory of Change, n.d.

Figure E.1
AM&E Decision Tree for Three Example Activities



SOURCE: RAND analysis.

RAND RR2165A-E.1

How Do I Develop a TOC?

The best way to develop a TOC is to work backward, starting with the objectives the activity is trying to achieve. The objectives should be mapped back to the outcomes leading to those objectives, the outputs that translate into specific outcomes, the activities linked to the outputs, and finally the inputs that are required to perform the activity.² This process creates a logical chain that underpins the activity. A TOC also makes explicit the often unstated causal assumptions for how activities achieve their intended outcomes.

Annex B: Recurring Activities

What Is a Recurring Activity?

A recurring activity is one that is likely to be carried out again in the future in a relatively similar way. To determine whether an activity is recurring, the planner should assess the activity across two dimensions. First, will the activity be repeated over time? If the activity is a one-off, then it is nonrecurring. Second, is the structure or design of the activity constant over time? If the way the activity is structured or designed changes, then the activity is nonrecurring.

For instance, most U.S. Army Training and Doctrine Command (TRADOC) courses are recurring activities because they are implemented every year with similar design and participants, and the topic for a particular TRADOC training remains stable over time. If the course structure were to change each year, then the activity would not be considered recurring.

Why Does It Matter Whether an Activity Is Recurring?

The opportunity to learn and improve is greater with a recurring activity because it is easier to understand whether and how the activity achieves its intended outcomes. In addition, the value of learning increases because the lessons learned from M&E can benefit the activity in the future, increasing effectiveness. Recurring activities generally allow and require more robust M&E processes.

An exception—an activity that is not strictly *recurring* but still warrants more robust AM&E—would be a pilot activity or a one-off activity that has an option for expansion. Pilots represent a good opportunity for evaluation purposes. If the activity provides valuable results, then it may be worth expanding. If it does not, then planners can consider changing the structure, topic, and other aspects of the activity before expansion.

Annex C: Resource Intensiveness

What Is Resource Intensiveness?

Resource-intensive activities require significant amounts of monetary or human resources to implement relative to other security cooperation activities. Planners should take a broad view of the “cost” of an activity, which should include personnel involved, time invested, direct monetary costs, and equipment utilized. An activity can be categorized as resource intensive if it requires significant resources in at least one of those dimensions.

² Center for Theory of Change, n.d.

In general, the planner should consider the accumulated resources that the activity requires within a fiscal year. An activity may involve limited resources each time it is implemented, but if it is implemented frequently then annual resource investments can be significant. For instance, traveling contact teams (TCTs) are integrated by U.S. military and civilian professionals who provide expertise to a host nation in a specific functional area. The duration of each of these activities tends to be short (e.g., five days), and the amount of monetary and human resources tends to be small. However, if TCTs were to happen frequently throughout the year they could add up to a significant amount of human and monetary resources. Alternatively, the activity may be conducted only sporadically but involve a significant amount of resources each time—thus making the total amount of resources required within a year substantial. Eager Light—a bilateral exercise with Jordan that occurs every year—is an example of a relatively infrequent but large exercise that is resource intensive.

There is no threshold for whether an activity is *resource intensive*. However, some activities are clearly relatively more costly. For instance, Multi-National Field Exercises, such as Eager Light (CPX), have a cost that is more than ten times that of implementing border patrol activities SMEE. This means that even nonrecurring activities may warrant additional AM&E.

Annex D: Priority Activities

What Determines Whether an Activity Is a Priority?

Priority activities are those that the Army considers relatively more important than others, and therefore their implementation justifies more in-depth AM&E. Activities that are recurring or resource intensive tend to be priorities, but nonrecurring and non-resource-intensive activities can also be a priority if, for example, they are mission critical, involve key stakeholders, or are activities the Army is considering expanding.

For instance, a humanitarian assistance activity could be one-off and relatively low cost for the Army across multiple resource dimensions. However, helping a particular country during a humanitarian crisis could also constitute a priority, and the Army may want to measure the results from that security cooperation investment. In this example a more robust AM&E approach may be warranted.

Annex E: AARs to Support AM&E

What Is the Value of an AAR in the Context of AM&E?

AARs provide a straightforward way for units to record and share basic information about how an activity is carried out and about its perceived effectiveness. An AAR is typically not a substitute for more systematic monitoring or information-gathering about security cooperation activities, but if used effectively AARs can support or supplement AM&E.

Objective and rigorous AARs are useful for AM&E across at least five dimensions:

- They provide data that help in monitoring, which can be valuable in improving the Army's capabilities and future performance at the tactical level.
- They provide data that enable understanding of whether an activity's implementation is aligned with broader strategic goals.

- They help the Army determine how its efforts can best support a partner nation's capabilities by providing information on the partner nation's needs.
- They increase knowledge-sharing capability within the Army, which supports learning.
- If AARs are part of an evaluation process, the Army can use them to systematically measure performance over time. This process can become a feedback loop in which actions directly inform all elements of the TOC and create more rigorous causal linkages between activities and outcomes, allowing the Army to showcase results for internal and external purposes.

How Can Planners Maximize the Value of AARs for AM&E?

For AARs to have maximal value for AM&E, the unit completing the AAR needs to (1) understand why the AAR has value for AM&E, (2) understand how the AAR will be used, and (3) receive clear guidance on what type of information they should include in the AAR *for the purpose of AM&E*. In addition, the AM&E value of AARs comes from planners' and executives' ability to access past AARs and/or summaries of past AARs that are relevant to a planned security cooperation activity. Security cooperation planners can help by synthesizing AARs and ensuring that relevant AM&E information is accessible to units.

Annex F: Monitoring

What Is Monitoring and What Purpose Does It Serve?

Monitoring is a systematic and recurring practice of collecting and analyzing information on strategic indicators while an activity is being implemented.³ Monitoring can be used to inform short-term decisions about the activity, compare the activity against targets, and correct the course of action to better align the activity with its strategic goals.⁴ Monitoring also tracks inputs, activities, outputs, and outcomes, but it does not try to establish whether the effects of the activity are *caused* by the activity; it does not establish causality. Monitoring can help answer questions such as:⁵

- How is the activity operating (e.g., is it efficient)?
- Are designated targets being met?
- Are expected outputs or outcomes likely to be produced?
- Who is implementing or delivering the different activity components?
- Who benefits from the activity?
- To what extent are activity beneficiaries satisfied with the services received?

How Are Activities Monitored?

Monitoring involves collecting data on indicators of inputs, activities, outputs (MOPs), and outcomes (MOEs).⁶ Indicators are usually numeric, even if they assess subjective components

³ Independent Evaluation Group, 2015.

⁴ Gertler et al., 2011.

⁵ Burt et al., 1997.

⁶ Gertler et al., 2011.

of an activity: they usually consist of counts, ratios, percentages, and so forth. However, monitoring does not need to be quantitative. It can be qualitative, but it must be carefully structured and systematic.

Indicators should follow the SMART criteria, which means they should be specific, measurable, attributable, realistic/relevant, and time-bound.⁷ Examples of output indicators for TRADOC activities are the number of training activities provided and the number of people attending the training. Examples of outcome indicators are the percentage of people trained who are actually using the training in their day-to-day jobs, and performance changes by those who received training. Monitoring an activity typically involves a systematic planning process that determines when and how data will be collected and who will collect the data.

Who Should Plan and Carry Out Monitoring?

The planner is usually responsible for planning data collection for indicators and for reporting on the data collected. Actual data collection should be carried out by an entity that has the appropriate capacity, skills, and access, which could be the planner, a unit, or in some cases a third party. The monitoring team should receive adequate training on how to conduct the activity-specific monitoring, and the planner should supervise the monitoring process.

How Much Does Monitoring Cost?

The cost of monitoring depends on the size of the activity, the number of indicators being tracked, the length of the activity, and the associated monitoring process. These costs can vary significantly. Monitoring can rely on either administrative data that are already being collected, which is a more basic type of monitoring exercise, or data collected through additional dedicated collection processes. Dedicated data collection is typically more costly, but it may be necessary to collect the right type of information to support AM&E needs.

Annex G: Evidence of Outcomes and Impacts

What Are Outcomes and Impacts?

Outcomes are the near-term results of an activity once it is completed. Impacts refer to the medium- or long-term changes due to the activity; they are also known as intermediate military objectives—goals the commands expect to achieve through their campaigns.⁸ Impacts take longer to be accomplished and can be influenced by multiple factors, including some that are outside the control of the planner. For these reasons, rigorous methodologies that separate the changes in *outcomes* due solely to the security cooperation activity from changes due to other factors that may also affect the outcomes are important for determining *impact*.

For instance, for activities associated with MEDCAP, planners are not able to control all outcomes, since the planner and unit cannot influence what beneficiaries do with the training once they leave the training course. An outcome would be the increased physician capacity after the training takes place, while an impact would be improving the partner nation's preventive health care systems.

⁷ Department of Defense, n.d.; Operational Policy and Quality Department, 2013.

⁸ Department of the Army, 2013b.

What Kind of Evidence Is Relevant to Establish a Causal Effect?

Evidence that establishes the causal effect an activity has on outcomes or impacts refers to robust and credible evidence about an activity's performance and whether the activity has accomplished the expected outcomes.⁹ Evidence should establish a causal effect and it should be rigorously measured. The strongest evidence will come from AM&E approaches that rely on a control or comparison group, such as impact evaluation methods. Impact evaluation allows the planner to estimate what would have happened to participants if they had not participated in the activity.

High-quality evidence does not need to be quantitative, but it often is. Moreover, using qualitative and quantitative methods together can produce robust impact estimates. However, evidence that is based exclusively on anecdotes or less rigorous quantitative analysis is not sufficient to establish the causal effect of an activity.

What Constitutes Sufficient and Clear Evidence?

There is no simple definition for sufficient evidence. One impact evaluation of an activity in a specific, narrow setting may not constitute sufficient evidence for an activity that is implemented in multiple settings and that may have multiple variations. Generally, evidence is strongest when it comes from multiple independent sources that use a range of high-quality methods. AM&E that is conducted without a comparison or control group will generally *not* provide sufficient evidence to establish that an activity causally achieves its outcomes and desired impacts.

Annex H: Performance Evaluation

What Is a Performance Evaluation and Why Is It Valuable?

Performance evaluation examines the degree to which the activity operations are working as intended, how the activity is implemented, and what the activity delivers.¹⁰ Performance evaluation is different from monitoring, which focuses on collecting information on an ongoing basis. Performance evaluations are conducted at discrete points in time, usually in the middle or at the end of an activity.¹¹ Performance evaluations do not provide causal evidence about outcomes or impacts of an activity; instead, they provide information on *implementation*.

For example, a performance evaluation could examine whether courses taught by the Army Training and Doctrine Command (i.e., military and professional training) are aligned with the activity objectives, whether the courses and associated resources are reaching the intended audience(s), and whether they are managed efficiently and effectively. Performance evaluations are more about *processes* than they are about *results*.

Even though performance evaluations are not able to establish whether the outcomes of interest are *caused* solely by the activity or by something else, process evaluations still provide valuable information that allows program planners to improve activity design, execution, and

⁹ Gertler et al., 2011.

¹⁰ J-PAL, 2016; Paul, 2013.

¹¹ Operations Policy and Quality Department, 2013.

delivery.¹² Process evaluations measure progress toward achieving expected targets, and they can support corrective action to improve an activity.¹³

How Much Does Performance Evaluation Cost?

Performance evaluation cost can vary considerably. They can often be implemented quickly by relying on administrative or monitoring data that have already been collected, which decreases the costs. However, there are some factors that need to be taken into account and that might increase costs. For instance, surveys and studies are usually conducted in process evaluations to complement existing data, and the larger the size and duration of the activity being evaluated, the larger the cost can become.

Annex I: Impact Evaluation

What Is an Impact Evaluation and Why Is It Valuable?

Impact evaluation measures the effect of an activity on an outcome of interest. For instance, the Army might be interested in assessing the impact of the MEDCAP program, which provides training to local physicians, training materials, and medications to the communities where it gets implemented. The Army might be interested in measuring the increase in the number of healthy people due to MEDCAP, and the improvements in the partner nation's preventive health care system achieved in the longer term. This type of measurement is targeted through impact evaluation.

The key to impact evaluation is measuring activity effectiveness by using methods that compare outcomes of participants with how these participants would have been had they not participated in the activity. Observing participants in both states (as participants and non-participants) is impossible; therefore, impact evaluation uses different methods to construct a group of nonparticipants that closely resembles the group of participants, but where the only difference between both groups is whether they participated in the activity.

Impact evaluation is different from other methods (e.g., comparing participant outcomes before and after the activity) because it allows attribution, which allows security cooperation planners to ensure that any outcomes—positive, negative, null, expected or unexpected—are caused by the activity and not by something else.¹⁴ Impact evaluation is also useful for policy-makers who have a mandate for accountability or who would like to understand whether an activity actually works (i.e., when there are evidence gaps), why it works, and by what magnitude. Impact evaluation also enables the identification of side effects and who benefits and who is harmed; and finally, it allows for comparing the effectiveness versus the cost of implementation of different activities with similar objectives.¹⁵ In the MEDCAP example, impact evaluation can determine whether MEDCAP has had an impact on outcomes such as decreased fatalities or illnesses in local populations, and estimates the magnitude of these changes.

¹² J-PAL, 2016.

¹³ World Bank, 2004.

¹⁴ World Bank, 2004.

¹⁵ J-PAL, 2016.

What Does an Impact Evaluation Cost?

The cost of an impact evaluation varies depending on the length of activity implementation, the variations in activity design, the number of participants and scope of the activity, the level of rigor of the evaluation, the expected time to observe an impact, and so forth. The cost can be as little as \$10,000 or as great as \$1 million. Both quasi-experimental and experimental approaches can be expensive, and their cost is driven mainly by data collection efforts.¹⁶

How Do I Determine Whether an Impact Evaluation Is Feasible?

Planners should analyze whether there is a lack of evidence on the effect of the activity on expected outcomes, whether the activity is recurring and resource intensive, whether the questions of interest cannot be answered through monitoring or performance evaluation, whether activities are stable (i.e., design and execution do not change significantly each time the activity is implemented), and whether there are enough resources (e.g., expertise, time) available to guarantee an adequate impact evaluation implementation. Since impact evaluations, and randomized evaluations in particular, can become very costly, activity planners should also assess whether the evidence provided by an impact evaluation is valuable enough to justify the costs involved.

If an Impact Evaluation Is Not Feasible, Are There Alternatives?

When impact evaluation is not possible, monitoring and performance evaluation can be used instead to provide useful information about whether the activity is being executed as planned, whether the expected outputs are delivered as expected, and whether activities are reaching target participants.¹⁷ Qualitative data are another alternative, as they can help with understanding the behaviors of participants, the level of satisfaction or dissatisfaction with the activity, and the environment that affects the outcomes of interest.

Qualitative data, monitoring, and performance evaluation should be seen as complements rather than substitutes in AM&E. However, when impact evaluation is not feasible, qualitative methods, monitoring, and performance evaluation can provide valuable insights for activity assessment.

Annex J: Consistency with Activity-Level TOC

What Does It Mean for an Activity to Be “Consistent with the TOC”?

For an activity to be consistent with the TOC, there should be consistency between the expected and actual inputs, outputs (MOPs), activities, and outcomes (MOEs). This comparison can be drawn from collecting data through monitoring and performance evaluation, which would answer questions such as the following: What are the activity inputs? Is the activity accomplishing the desired outputs? Are the outputs linked to the desired outcomes? What are the short- and medium-term outcomes? Is the activity being implemented as planned?

¹⁶ Shah et al., 2015.

¹⁷ J-PAL, *Deciding What, When, and Where to Invest in Impact Evaluation*, n.d.

What Else Is Relevant for Assessing Activity Performance?

A TOC is static and based on stated and unstated assumptions that are required to maintain the connections between inputs, activities, outputs, and outcomes. While assessing activity performance, or consistency with the TOC, evaluators should consider not only the assumptions and linkages established in the TOC but the activity's unintended consequences (i.e., side effects), the availability of inputs, the enrollment and attrition rates, and other contextual factors that may affect performance.

Annex K: Objectives and Setting

How Do We Know If an Activity's Objectives Have Changed and Why These Changes Matter?

The effectiveness of an activity, and its consistency with the TOC, may not hold if the activity objective changes. The Army measures activity effectiveness and consistency with the TOC for a particular objective. If the activity objective changes, then the Army may have to conduct more in-depth results measurement to establish whether the activity, under the new objective, is still effective and aligned with the overall TOC (i.e., following the established inputs, outputs, and outcomes).

For instance, consider the hypothetical case in which the Army has established that a TCT for chemical-biological defense is effective and consistent with the activity-level TOC. If the Army wants to establish this same relationship for a TCT on cooperative threat reduction, then monitoring may not be enough because the topic of the TCT has changed, and the Army may also consider process evaluation (as shown in boxes 18 and 19 of the decision tree).

What Type of a Change in Setting Is Relevant and Why Does This Matter for AM&E?

The results obtained from evaluating an activity in a particular setting may not hold when an activity is planned for a different setting. The change in the environment affects the way in which inputs translate into outputs, and outputs translate into outcomes. For instance, a humanitarian activity implemented in the Middle East that performed successfully may be unsuccessful when implemented in East Asia. The weather, local diseases, and geography in East Asia may prove more challenging than in the Middle East for planners. As a result, the activity planner should consider conducting further evaluations to assess whether activity effects hold (i.e., are externally valid) regardless of the environment.

Works Included in Literature Review

Table F.1
Works Included in Literature Review

Mediating Factor	Authors	Title	Publication Date	Type of Material	Method Used
Absorptive capacity	Rand, Dafna H., and Stephen Tankel	Security Cooperation & Assistance: Rethinking the Return on Investment	2015	Report	Mixed methods
Absorptive capacity	Lamb, Robert D., and Kathryn Mixon	Rethinking Absorptive Capacity: A New Framework, Applied to Afghanistan's Police Training Program	2013	Report	Qualitative
Aid inertia	Berger, L.	Guns, Butter, and Human Rights: The Congressional Politics of U.S. Aid to Egypt	2012	Journal article	Quantitative
Country stability or fractionalization	Creasey, Ellyn, Ahmed S. Rahman, and Katherine A. Smith	Does Nation Building Spur Economic Growth?	2015	Journal article	Quantitative
Country stability or fractionalization	Jones, Seth, et al.	Securing Tyrants or Fostering Reform? U.S. Internal Security Assistance to Repressive and Transitioning Regimes	2006	Report	Qualitative
Country stability or fractionalization	McNerney, Michael, et al.	Assessing Security Cooperation as a Preventive Tool	2014	Report	Mixed methods
Country stability or fractionalization	Mehler, Andreas	Why Security Forces Do Not Deliver Security: Evidence from Liberia and the Central African Republic	2012	Journal article	Qualitative
Country stability or fractionalization	Ruby, Tomislav Z., and Douglas Gibler	US Professional Military Education and Democratization Abroad	2010	Journal article	Mixed methods
Country stability or fractionalization	Maniruzzaman, Talukder	Arms Transfers, Military Coups, and Military Rule in Developing States	1992	Journal article	Mixed methods

Table F.1—Continued

Mediating Factor	Authors	Title	Publication Date	Type of Material	Method Used
Economic policies	DeRouen, K., and U. Heo	Modernization and the Military in Latin America	2001	Journal article	Quantitative
Economic policies	Poe, S.C., and J. Meernik	US Military Aid in the 1980s—a Global Analysis	1995	Journal article	Quantitative
Economic policies	Demirel-Pegg, Tejin, and James Moskowitz	US Aid Allocation: The Nexus of Human Rights, Democracy, and Development	2009	Journal article	Quantitative
Economic policies	Creasey, Ellyn, Ahmed S. Rahman, and Katherine A. Smith	Does Nation Building Spur Economic Growth?	2015	Journal article	Quantitative
History/common language and/or cultural awareness	Alesina, Alberto, and David Dollar	Who Gives Foreign Aid to Whom and Why?	2000	Journal article	Quantitative
History/common language and/or cultural awareness	O'Mahony et al.	Assessing the Value of Regionally Aligned Forces in Army Security Cooperation	2017	Report	Mixed methods
History/common language and/or cultural awareness	Carter, David B. and Randall W. Stone	Democracy and Multilateralism: The Case of Vote Buying in the UN General Assembly	2013	Working paper	Game theoretic model
History/common language and/or cultural awareness	Abbe, Allison	Building Cultural Capability for Full-Spectrum Operations	2008	Thesis	Qualitative
History/common language and/or cultural awareness	Taw, Jennifer Morrison, and William H. McCoy	International Military Student Training: Beyond Tactics	1992	Report	Qualitative
Institutional capacity	Crawford, Gordon	Foreign Aid and Political Conditionality: Issues of Effectiveness and Consistency	1997	Journal article	Qualitative
Institutional capacity	Jones, Seth, et al.	Securing Tyrants or Fostering Reform? U.S. Internal Security Assistance to Repressive and Transitioning Regimes	2006	Report	Qualitative
Institutional capacity	Mehler, Andreas	Why Security Forces Do Not Deliver Security: Evidence from Liberia and the Central African Republic	2012	Journal article	Qualitative
Institutional capacity	Atkinson, Carol	Constructivist Implications of Material Power: Military Engagement and the Socialization of States, 1972–2000	2006	Journal article	Quantitative
Noneconomic policies	Poe, S. C., and J. Meernik	US Military Aid in the 1980s—a Global Analysis	1995	Journal article	Quantitative
Noneconomic policies	Poe, Steven C.	Human Rights and the Allocation of U.S. Military Assistance	1991	Journal article	Quantitative

Table F.1—Continued

Mediating Factor	Authors	Title	Publication Date	Type of Material	Method Used
Noneconomic policies	Carter, David B., and Randall W. Stone	Democracy and Multilateralism: The Case of Vote Buying in the UN General Assembly	2013	Working paper	Game theoretic model
Noneconomic policies	Demirel-Pegg, Tejin, and James Moskowitz	US Aid Allocation: The Nexus of Human Rights, Democracy, and Development	2009	Journal article	Quantitative
Noneconomic policies	Blanton, Shannon Lindsey	Promoting Human Rights and Democracy in the Developing World: Rhetoric versus U.S. Arms Exports	2000	Journal article	Quantitative
Noneconomic policies	Blanton, Shannon Lindsey	Foreign Policy in Transition? Human Rights, Democracy, and U.S. Exports	2005	Journal article	Quantitative
Noneconomic policies	Cingranelli, David L, and Thomas E. Pasquarello	Human Rights Practices and the Distribution of U.S. Foreign Aid to Latin American Countries	1985	Journal article	Quantitative
Noneconomic policies	Carleton, David B., and Randall W. Stohl	The Role of Human Rights in U.S. Foreign Assistance Policy: A Critique and Reappraisal	1987	Journal article	Quantitative
Noneconomic policies	McCormick, James M., and Neil Mitchell	Is U.S. Aid Really Linked to Human Rights in Latin America?	1988	Journal article	Quantitative
Noneconomic policies	Crawford, Gordon	Foreign Aid and Political Conditionality: Issues of Effectiveness and Consistency	1997	Journal article	Qualitative
Noneconomic policies	Carey, Sabine C.	European Aid: Human Rights versus Bureaucratic Inertia?	2007	Journal article	Quantitative
Noneconomic policies	Mehler, Andreas	Why Security Forces Do Not Deliver Security: Evidence from Liberia and the Central African Republic	2012	Journal article	Qualitative
Noneconomic policies	Taw, Jennifer Morrison, and William H. McCoy	International Military Student Training: Beyond Tactics	1992	Report	Qualitative
Noneconomic policies	Whitaker, Beth Elise	Compliance Among Weak States: Africa and the Counter-Terrorism Regime	2010	Journal article	Qualitative
Noneconomic policies	Sislin, John	Arms as Influence: The Determinants of Successful Influence	1994	Journal article	Quantitative
Noneconomic policies	Atkinson, Carol	Constructivist Implications of Material Power: Military Engagement and the Socialization of States, 1972–2000	2006	Journal article	Quantitative

Table F.1—Continued

Mediating Factor	Authors	Title	Publication Date	Type of Material	Method Used
Noneconomic policies	Apodaca, Clair, and Michael Stohl	United States Human Rights Policy and Foreign Assistance	1999	Journal article	Quantitative
Quantity of aid	Paul, Christopher, et al.	What Works Best When Building Partner Capacity, and Under What Circumstances?	2013	Report	Mixed methods
Quantity of aid	Moroney, Jennifer, et al.	Review of Security Cooperation Mechanisms Combatant Commands Utilize to Build Partner Capacity	2013	Report	Qualitative
Quantity of aid	Whitaker, Beth Elise	Compliance Among Weak States: Africa and the Counter-Terrorism Regime	2010	Journal article	Qualitative
Quantity of aid	Sislin, John	Arms as Influence: The Determinants of Successful Influence	1994	Journal article	Quantitative
Specific project characteristics	Moroney, Jennifer, et al.	Review of Security Cooperation Mechanisms Combatant Commands Utilize to Build Partner Capacity	2013	Report	Qualitative
Specific project characteristics	International Security Advisory Board	Report on Security Capacity Building	2013	Report	Qualitative
Specific project characteristics	Government Accountability Office	Key Practices to Effectively Manage Department of Defense Efforts to Promote Security Cooperation	2013	Report	Qualitative
Specific project characteristics	Government Accountability Office	Project Evaluations and Better Information Sharing Needed to Manage the Military's Efforts	2012	Report	Mixed methods
Specific project characteristics	Moroney, Jennifer D. P., et al.	How Successful are U.S. Efforts to Build Capacity in Developing Countries? A Framework to Assess the Global Train and Equip "1206" Program	2011	Report	Qualitative
Spillover effects	Maniruzzaman, Talukder	Arms Transfers, Military Coups, and Military Rule in Developing States	1992	Journal article	Mixed methods
Spillover effects	Blanton, Shannon Lindsey	Instruments of Security or Tools of Repression? Arms Imports and Human Rights Conditions in Developing Countries	1999	Journal article	Quantitative
Strategic goals and relationships	Alesina, Alberto, and David Dollar	Who Gives Foreign Aid to Whom and Why?	2000	Journal article	Quantitative

Table F.1—Continued

Mediating Factor	Authors	Title	Publication Date	Type of Material	Method Used
Strategic goals and relationships	Berger, L.	Guns, Butter, and Human Rights—the Congressional Politics of U.S. Aid to Egypt	2012	Journal article	Quantitative
Strategic goals and relationships	Bapat, N. A.	Transnational Terrorism, U.S. Military Aid, and the Incentive to Misrepresent	2011	Journal article	Quantitative
Strategic goals and relationships	Paul, Christopher, et al.	What Works Best When Building Partner Capacity, and Under What Circumstances?	2013	Report	Mixed methods
Strategic goals and relationships	O'Mahony et al.	Assessing the Value of Regionally Aligned Forces in Army Security Cooperation	2017	Report	Mixed methods
Strategic goals and relationships	Meernik, Kruger, and Poe	Testing Models of U.S. Foreign Policy: Foreign Aid During and After the Cold War	1998	Journal article	Quantitative
Strategic goals and relationships	Congressional Research Services	What Is “Building Partner Capacity?” Issues for Congress	2015	Report	Qualitative
Strategic goals and relationships	Whitaker, Beth Elise	Compliance Among Weak States: Africa and the Counter-Terrorism Regime	2010	Journal article	Qualitative
Time periods and horizons	Bapat, N. A.	Transnational Terrorism, U.S. Military Aid, and the Incentive to Misrepresent	2011	Journal article	Quantitative
Time periods and horizons	Paul, Christopher, et al.	What Works Best When Building Partner Capacity, and Under What Circumstances?	2013	Report	Mixed methods
Time periods and horizons	Moroney, Jennifer, et al.	Review of Security Cooperation Mechanisms Combatant Commands Utilize to Build Partner Capacity	2013	Report	Qualitative
Type of aid	Sislin, John	Arms as Influence: The Determinants of Successful Influence	1994	Journal article	Quantitative
Type of aid	Neumayer, Eric	Is Respect for Human Rights Rewarded? An Analysis of Total Bilateral and Multilateral Aid Flows	2003	Journal article	Quantitative

Further Reading Suggestions on AM&E and Impact Evaluation Practices

General Concepts, Methods, and Implementation

Bamberger, Michael, et al., *Real World Evaluation: Working Under Budget, Time, Data, and Political Constraints*, Thousand Oaks, Calif.: SAGE Publications, 2013.

Gertler, Paul J., et al., *Impact Evaluation in Practice*, 2nd ed., Washington, D.C.: World Bank Publications, 2016. As of January 3, 2017:

<http://www.worldbank.org/en/programs/sief-trust-fund/publication/impact-evaluation-in-practice>

Khandker, Shahidur R., et al., *Handbook on Impact Evaluation: Quantitative Methods and Practices*, Washington, D.C.: World Bank Publications, 2009. As of January 3, 2017:

<https://openknowledge.worldbank.org/bitstream/handle/10986/2693/520990PUB0EPI1101Official0Use0Only1.pdf>

Prennushi, G., G. Rubio, and K. Subbarao, "Monitoring and Evaluation," in *Sourcebook for Poverty Reduction Strategies*, Washington, D.C.: World Bank Publications, 2000. As of January 3, 2017:

http://siteresources.worldbank.org/INTPRS1/Resources/383606-1205334112622/4480_chap3.pdf

Shadish, William, Thomas D. Cook, and Donald T. Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Boston: Houghton Mifflin, 2002.

AM&E Framework Development and Implementation

Bamberger, Michael, et al., *Real World Evaluation: Working Under Budget, Time, Data, and Political Constraints*, Thousand Oaks, Calif.: SAGE Publications, 2011.

Development Assistance Committee, *Principles for Evaluation of Development Assistance*, Paris: OECD, 1991. As of January 3, 2017:

<http://www.oecd.org/dac/evaluation/50584880.pdf>

Keith, Mackay, *How to Build M&E Systems to Support Better Government*, Washington, D.C.: World Bank Publications, 2007. As of January 3, 2017:

http://siteresources.worldbank.org/EXTEVACAPDEV/Resources/4585672-1251737367684/How_to_build_ME_gov.pdf

Menon, Saraswathi, Judith Karl, and Kanni Wignaraja, *Handbook on Planning, Monitoring and Evaluating for Development Results*, New York: UNDP Evaluation Office, 2009. As of January 3, 2017:

<http://web.undp.org/evaluation/evaluations/handbook/english/documents/pme-handbook.pdf>

Roberts, Dawn, and Nidhi Khattri, *Designing a Results Framework for Achieving Results: A How-To Guide*, Washington, D.C.: World Bank Publications, 2012. As of January 3, 2017:

http://siteresources.worldbank.org/EXTEVACAPDEV/Resources/designing_results_framework.pdf

USAID, "USAID Evaluation Toolkit: Guidance, Tools, & Resources for Planning, Managing, & Learning from Evaluations," 2016. As of January 3, 2017:

<https://usaidlearninglab.org/evaluation?tab=1>

Security Cooperation or Army Practices

Department of Defense, “Assessment, Monitoring, and Evaluation Policy for the Security Cooperation Enterprise.” As of January 3, 2017:
<http://www.dtic.mil/whs/directives>

Marquis, Jeff P., et al., *Developing an Assessment, Monitoring, and Evaluation Framework for U.S. Department of Defense Security Cooperation*, Santa Monica, Calif.: RAND Corporation, 2016. As of January 3, 2017:
http://www.rand.org/pubs/research_reports/RR1611.html

McNerney, Michael J., et al., *SMART Security Cooperation Objectives: Improving DoD Planning and Guidance*, Santa Monica, Calif.: RAND Corporation, 2016. As of January 3, 2017:
http://www.rand.org/pubs/research_reports/RR1430.html

Paul, Christopher, “Foundations and Assessment: The Hierarchy of Evaluation and the Importance of Articulating a Theory of Change,” *Small Wars Journal*, July 30, 2013.

Theories of Change

Center for Theory of Change, “Theories of Change,” n.d. As of January 3, 2017:
<http://www.theoryofchange.org>

Vogel, Isabel, *Review of the Use of ‘Theory of Change’ in International Development*, UK Department of International Development, 2012. As of January 3, 2017:
http://www.theoryofchange.org/pdf/DFID_ToC_Review_VogelV7.pdf

Qualitative Methods

Quinn Patton, Michael, *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*, 4th ed., Thousand Oaks, Calif.: SAGE Publications, 2015.

Randomized Experiments

Banerjee, Abhijit, and Esther Duflo, *Handbook of Field Experiments*, Amsterdam: North Holland Publications, 2017. As of January 3, 2017:
<https://www.povertyactionlab.org/handbook-field-experiments>

Gerber, Alan S., and Donald P. Green, *Field Experiments: Design, Analysis, and Interpretation*, New York: W. W. Norton, 2012.

Glennerster, Rachel, and Kudzai Takavarasha, *Running Randomized Evaluations: A Practical Guide*, Princeton, N.J.: Princeton University Press, 2013.

Karlan, D., and J. Appel, *Failing in the Field: What We Can Learn When Field Research Goes Wrong*, Princeton, N.J.: Princeton University Press, 2016.

MIT at EdX, “Evaluating Social Programs: Understand How Randomized Evaluations Can Be Used to Evaluate Social and Development Programs,” n.d. As of January 3, 2017:
<https://www.edx.org/course/evaluating-social-programs-mitx-jpal101x-4#!>

References

- Abbe, Allison, *Building Cultural Capability for Full-Spectrum Operations*, Arlington, Va.: United States Army Research Institute for the Behavioral and Social Sciences, 2008.
- Abbe, Allison, Lisa M. V. Gulick, and Jeffrey L. Herman, *Cross-Cultural Competence in Army Leaders: A Conceptual and Empirical Foundation*, Arlington, Va.: United States Army Research Institute for the Behavioral and Social Sciences, 2007.
- Angrist, Joshua, and Jorn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, N.J.: Princeton University Press, 2009.
- Apodaca, Clair, and Michael Stohl, "United States Human Rights Policy and Foreign Assistance," *International Studies Quarterly*, Vol. 43, No. 1, 1999, pp. 185–198.
- Army Capabilities Integration Center, *Army Warfighting Challenges*, February 27, 2017. As of March 10, 2017: <http://www.arcic.army.mil/Initiatives/ArmyWarfightingChallenges>
- Atkinson, Carol, "Constructivist Implications of Material Power: Military Engagement and the Socialization of States, 1972–2000," *International Studies Quarterly*, Vol. 50, No. 3, 2006, pp. 509–537.
- Bamberger, Michael, "Introduction to Mixed Methods in Impact Evaluation," *InterAction*, Impact Evaluation Notes, No. 3, August 2012, pp. 1–38. As of January 3, 2017: [https://www.interaction.org/sites/default/files/Mixed%20Methods%20in%20Impact%20Evaluation%20\(English\).pdf](https://www.interaction.org/sites/default/files/Mixed%20Methods%20in%20Impact%20Evaluation%20(English).pdf)
- Bamberger, Michael, Jim Rugh, and Linda S. Mabry, *Real World Evaluation: Working Under Budget, Time, Data, and Political Constraints*, Thousand Oaks, Calif.: SAGE Publications, 2013.
- Bapat, Navin A., "Transnational Terrorism, US Military Aid, and the Incentive to Misrepresent," *Journal of Peace Research*, Vol. 48, No. 3, 2011, pp. 303–318.
- Behaghel, Luc, and Philippe Zamora, "Randomized Controlled Trial: A Historical Perspective," *Abdul Latif Jameel Poverty Action Lab*, J-PAL Advanced Course, Paris, 2012. As of January 3, 2017: <https://www.povertyactionlab.org/sites/default/files/L1-historical-intro.pdf>
- Berger, Lars, "Guns, Butter, and Human Rights: The Congressional Politics of U.S. Aid to Egypt," *American Politics Research*, Vol. 40, No. 4, 2012, pp. 603–635.
- Biddle, Stephen, Ryan Baker, and Julia MacDonald, "Small Footprint, Small Payoff: The Military Effectiveness of Security Force Assistance," Working Paper, September 8, 2015.
- Blair, Dennis, *Military Engagement: Influencing Armed Forces Worldwide to Support Democratic Transitions, Volume 1: Overview and Action Plan*, Washington, D.C.: Brookings Institution Press, 2013.
- Blanton, Shannon Lindsey, "Instruments of Security or Tools of Repression? Arms Imports and Human Rights Conditions in Developing Countries," *Journal of Peace Research*, Vol. 36, No. 2, 1999, pp. 233–244.
- Blomquist, John, "Impact Evaluation of Social Programs: A Policy Perspective," *World Bank*, Social Safety Nets Primer Notes, No. 14, September 2003. As of January 3, 2017: <https://openknowledge.worldbank.org/handle/10986/11827>

Bonbright, David, Fred Carden, Sarah Earl, Sanjeev Khagram, Nancy MacPherson, Zenda Ofir, and Patricia Rogers, *Impact Evaluation for Improving Development (ie4id)*, 2010. As of January 3, 2017: https://evaluationrevisited.files.wordpress.com/2010/08/ie4id_final_draft_04-14-10.pdf

Bosworth, Barry P., and Susan M. Collins, "The Empirics of Growth: An Update," *Brookings Papers on Economic Activity*, Vol. 2, 2003, pp. 113–206.

Brown, Julia, Lucia Goin, Nora Gregory, Katherine Hoffmann, and Kim Smith, *Evaluating Financial Products and Services in the US: A Toolkit for Running Randomized Controlled Trials*, New Haven, Conn.: Innovations for Poverty Action, 2015. As of January 3, 2017: [https://www.poverty-action.org/sites/default/files/publications/Evaluating%20Financial%20Products%20and%20Services%20in%20the%20US%20-%20A%20Toolkit%20for%20Running%20RCTs%20\(1\).pdf](https://www.poverty-action.org/sites/default/files/publications/Evaluating%20Financial%20Products%20and%20Services%20in%20the%20US%20-%20A%20Toolkit%20for%20Running%20RCTs%20(1).pdf)

Bryant, Benjamin P., and Robert J. Lempert, "Thinking Inside the Box: A Participatory Computer-Assisted Approach to Scenario Discovery," *Technological Forecasting and Social Change*, Vol. 77, No. 1, January 2010, pp. 34–49.

Burnside, Craig, and David Dollar, "Aid, Policies and Growth," *American Economic Review*, Vol. 90, No. 4, 2000, pp. 847–868.

———, "Aid, Policies, and Growth: Reply," *The American Economic Review*, Vol. 94, No. 3, 2004, pp. 781–784.

Burt, Martha R., Adele V. Harrell, Lisa C. Newmark, Landon Y. Aron, and Lisa K. Jacobs, "Chapter 6: Choosing an Evaluation Design," in *Evaluation Guidebook*, Washington, D.C.: Urban Institute, 1997. As of January 3, 2017: <https://www.bja.gov/evaluation/guide/documents/stop1-4.html>

Cameron, Drew, and Jorge Miranda, "Trends in Impact Evaluation: Did We Ever Learn?" *International Initiative for Impact Evaluation*, 2015. As of January 3, 2017: <http://blogs.3ieimpact.org/trends-in-impact-evaluation-did-we-ever-learn>

Carleton, David B., and Randall W. Stohl, "The Role of Human Rights in U.S. Foreign Assistance Policy: A Critique and Reappraisal," *American Journal of Political Science*, Vol. 31, No. 4, 1987, pp. 1002–1018.

Center for Army Lessons Learned, *Security Cooperation: Lessons and Best Practices*, Bulletin: 16-09, March 2016. As of January 3, 2017: <http://usacac.army.mil/sites/default/files/publications/16-09.pdf>

Center for Theory of Change, "What Is a Theory of Change," n.d. As of January 3, 2017: <http://www.theoryofchange.org/what-is-theory-of-change>

Cingranelli, David L., and Thomas E. Pasquarello, "Human Rights Practices and the Distribution of U.S. Foreign Aid to Latin American Countries," *American Journal of Political Sciences*, Vol. 29, No. 3, 1985, pp. 539–563.

Cingranelli, David, David L. Richards, and K. Chad Clay, "The CIRI Human Rights Dataset," *CIRI Human Rights Data Project*, Version 2014.04.14, 2014. As of January 3, 2017: <http://www.humanrightsdata.com/p/data-documentation.html>

Clemens, Michael A., Steven Radelet, Rikhil R. Bhavnani, and Samuel Bazzi, "Counting Chickens When They Hatch: The Short Term Effect of Aid on Growth," Center for Global Development, Working Paper 44, 2004.

———, "Counting Chickens When They Hatch: Timing and the Effects of Aid on Growth," *The Economic Journal*, Vol. 122, No. 561, 2012, pp. 590–617.

Collier, Paul, and Anke Hoeffler, "Aid, Policy and Growth in Post-Conflict Societies," *European Economic Review*, Vol. 48, No. 5, 2004, pp. 1125–1145.

Creasey, Ellyn, Ahmed S. Rahman, and Katherine A. Smith, "Does Nation Building Spur Economic Growth?" *Economic Inquiry*, Vol. 53, No. 1, January 2015, pp. 660–680.

"CSA Whitepaper on POTUS Strategy and Global Landpower Network: Delivering National Power on Land Through the GLN," July 23, 2014.

- Defense Security Cooperation Agency, "Foreign Military Financing (FMF)," n.d.a. As of January 3, 2017: <http://www.dsca.mil/programs/foreign-military-financing-fmf>
- , "Foreign Military Sales (FMS)," n.d.b. As of January 3, 2017: <http://www.dsca.mil/programs/foreign-military-sales-fms>
- , "International Military Education and Training (IMET)," n.d.c. As of January 3, 2017: <http://www.dsca.mil/programs/international-military-education-training-imet>
- DeNault, Tiffany, "Army Civil Affairs Provides Medical Care to Djibouti Village," May 21, 2016. As of January 3, 2017: https://www.army.mil/article/168358/army_civil_affairs_provides_medical_care_to_djibouti_village
- Department of the Army, *ADRP 5-0: The Operations Process*, Washington, D.C.: Army Doctrine Reference Publication Headquarters, May 17, 2012.
- , *FM 3-22: Army Support to Security Cooperation*, Washington, D.C.: Army Doctrine Reference Publication Headquarters, January 2013a.
- , *APD 11-31: Army Security Cooperation Handbook*, Washington, D.C.: Army Doctrine Reference Publication Headquarters, March 5, 2013b.
- , *Army Regulation 11-31: Army Security Cooperation Policy*, Washington, D.C.: Army Doctrine Reference Publication Headquarters, March 21, 2013c.
- , *Army Strategic Guidance for Security Cooperation: An Enduring Mission for "Prevent, Shape, Win,"* Washington, D.C.: Army Doctrine Reference Publication Headquarters, August 2014a.
- , *Global-Theater Security Cooperation Management System (G-TSCMIS) Guidance and Business Rules*, PR 2456: GLN report, Washington, D.C.: Army Doctrine Reference Publication Headquarters, 2014b.
- , *APD 11-31: Army Security Cooperation Handbook*, Washington, D.C.: Army Doctrine Reference Publication Headquarters, February 6, 2015.
- , *U.S. Army Security Cooperation Planners' Course Student Handbook, Deputy Chief of Staff, G-3/5/7, Strategy, Plans and Policy Directorate*, April 11–16, 2016.
- Department of Defense, "Assessment, Monitoring, and Evaluation Policy for the Security Cooperation Enterprise," n.d. As of January 3, 2017: <http://www.dtic.mil/whs/directives>
- , *Quadrennial Defense Review 2014*, Washington, D.C., 2014.
- Department for International Development, "DFID Research Strategy 2008–2013," n.d. As of January 3, 2017: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/67757/research-strategy-08.pdf
- DeRouen, Karl, and Uk Heo, "Modernization and the Military in Latin America," *British Journal of Political Science*, Vol. 31, No. 3, 2001, pp. 475–496.
- Dreher, Axel, Jan-Egbert Sturm, and James Raymond Vreeland, "Does Membership on the UN Security Council Influence IMF Decisions? Evidence from Panel Data," CESifo Working Paper Series No. 1808, August 2006.
- , "Development Aid and International Politics: Does Membership on the UN Security Council Influence World Bank Decisions?" *Journal of Development Economics*, Vol. 88, No.1, 2009, pp. 1–18.
- Development Impact Evaluation, *Transforming Development Through Impact Evaluation, Annual Report, March 2015 to March 2016*, World Bank, 2016. As of January 3, 2017: <http://pubdocs.worldbank.org/en/350311472677341646/i2i-Annual-Report-2016-8-31-16web.pdf>
- Duponchel, Marguerite, Lisa Chauvet, and Paul Collier, "What Explains Aid Project Success in Post-Conflict Situations?" World Bank Policy Research Working Paper Series, 2010.
- Dutta, Nabamita, Deepraj Mukherjee, and Sanjukta Roy, "Re-Examining the Relationship Between Domestic Investment and Foreign Aid: Does Political Stability Matter?" *International Review of Applied Economics*, Vol. 29, No. 3, 2015, pp. 259–286.

Easterly, William, and Ross Levine, "Africa's Growth Tragedy: Policies and Ethnic Divisions," *The Quarterly Journal of Economics*, Vol. 112, No. 4, 1997, pp. 1203–1250.

Easterly, William, Ross Levine, and David Roodman, "New Data, New Doubts: A Comment on Burnside and Dollar's 'Aid, Policies, and Growth' (2000)," National Bureau of Economic Research Working Paper No. 9846, July 2003.

Egel, Daniel, Charles P. Ries, Ben Connable, Todd C. Helmus, Eric Robinson, Isaac Baruffi, Melissa A. Bradley, Kurt Card, Kathleen Loa, Sean Mann, Fernando Sedano, Stephan B. Seabrook, and Robert Stewart, *Investing in the Fight: Assessing the Use of the Commander's Emergency Response Program in Afghanistan*, Santa Monica, Calif.: RAND Corporation, RR-1508-OSD, 2016. As of January 3, 2017: http://www.rand.org/pubs/research_reports/RR1508.html

Ezemenari, Kene, Anders Rudqvist, and K. Subbarao, *Impact Evaluation: A Note on Concepts and Methods*, World Bank Poverty Reduction and Economic Management Network, Washington, D.C.: World Bank, 1999.

Feeny, Simon, and Ashton de Silva, "Measuring Absorptive Capacity Constraints to Foreign Aid," *Economic Modeling*, Vol. 29, No. 3, 2012, pp. 725–733.

Friedman, Jerome H., and Nicholas I. Fisher, "Bump Hunting in High-Dimensional Data," *Statistics and Computing*, Vol. 9, No. 2, 1999, pp. 123–143.

Fuchs, Andreas, Axel Dreher, and Peter Nunnenkamp, "Determinants of Donor Generosity: A Survey of the Aid Budget Literature," *World Development*, Vol. 56, April 2014, pp. 172–199.

Gaarder, Marie, and Jeannie Annan, "Impact Evaluation of Conflict Prevention and Peacebuilding Interventions," World Bank Policy Research Working Paper #6496, June 2013.

"Garuda Shield: Exercise Garuda Shield," n.d. As of January 3, 2017: <https://www.facebook.com/ExerciseGarudaShield/>

Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeesch, *Impact Evaluation in Practice*, Washington, D.C.: Inter-American World Bank Publications, 2011. As of January 3, 2017: <https://openknowledge.worldbank.org/handle/10986/25030>

Glennerster, Rachel, "The Practicalities of Running Randomized Evaluations: Partnerships, Measurement, Ethics, and Transparency," in *Handbook of Field Experiments*, J-PAL, 2016. As of January 3, 2017: https://www.povertyactionlab.org/sites/default/files/publications/Practical%20issues%20in%20RCTs_Glennerster%20.pdf

Green, Duncan, "What Use Is a Theory of Change? 6 Benefits, and Some Things to Avoid," *People, Spaces and Deliberation*, World Bank, 2013. As of January 3, 2017: <http://blogs.worldbank.org/publicsphere/what-use-theory-change-6-benefits-and-some-things-avoid>

Gueron, Judith, "The Politics and Practice of Social Experiments: Seeds of a Revolution," in *Handbook of Field Experiments*, Amsterdam: North Holland Publishing Co., March 21, 2016. As of January 3, 2017: <https://www.povertyactionlab.org/sites/default/files/publications/GUERON%20POLITICS%20AND%20PRACTICE%20OF%20SOCIAL%20EXPERIMENTS%203.22.2016%20FINAL.pdf>

Gugerty, Mary Kay, Dean Karlan, and Delia Welsh, *Using Administrative Data for Monitoring and Evaluation*, Goldilocks Deep Dive, Innovations for Poverty Action, 2016. As of January 3, 2017: http://www.poverty-action.org/sites/default/files/publications/Goldilocks-Deep-Dive-Using-Administrative-Data-for-Monitoring-and-Evaluation_1.pdf

Headquarters, Department of the Army, *FM 3-22: Army Support to Security Cooperation*, Washington, D.C.: Department of the Army, January 2013.

Heidelberg Institute for International Conflict Research, *Conflict Barometer*, n.d. As of January 27, 2017: https://www.hiik.de/en/konfliktbarometer/pdf/ConflictBarometer_2015.pdf

IHS Jane's, *Sentinel Security Assessments*, 2015.

Independent Evaluation Group, "Approach Paper Report on Self-Evaluation Systems (ROSES 2016)," *World Bank Group*, No. 95189, February 24, 2015. As of January 3, 2017: <http://documents.worldbank.org/curated/en/481241468321544453/pdf/951890WP0ROSES000PUBLIC00Box391409B.pdf>

- Institute for Economics & Peace, *Global Peace Index, 2016: Ten Years of Measuring Peace*, Institute for Economics & Peace IEP Report 30, June 2016. As of March 10, 2017:
http://economicsandpeace.org/wp-content/uploads/2016/06/GPI-2016-Report_2.pdf
- International Security Advisory Board, *Report on Security Capacity Building*, Washington, D.C.: International Security Advisory Board, 2013.
- INTRAC, *Monitoring and Evaluation: New Developments and Challenges*, International Conference, Soesterberg, The Netherlands, June 14–16, 2011. As of January 3, 2017:
<https://www.intrac.org/wpcms/wp-content/uploads/2016/09/Monitoring-and-Evaluation-New-developments-and-challenges-conference-report.pdf>
- Joint Chiefs of Staff, *Joint Publication 3-20*, draft, 2016 (provided by sponsor in summer, 2016).
- Jones, Seth G., Olga Oliker, Peter Chalk, C. Christine Fair, Rollie Lal, and James Dobbins, *Securing Tyrants or Fostering Reform? U.S. Internal Security Assistance to Repressive and Transitioning Regimes*, Santa Monica, Calif.: RAND Corporation, MG-550-OSI, 2006. As of August 11, 2011:
<https://www.rand.org/pubs/monographs/MG550.html>
- J-PAL, *Deciding What, When, and Where to Invest in Impact Evaluation*, n.d. As of January 3, 2017:
https://www.povertyactionlab.org/sites/default/files/event/LI0_Discussion_RCTs_In_Broader_Strategy.pdf
- , “Introduction to Evaluations,” *Impact Evaluation*, 2016. As of January 3, 2017:
<https://www.povertyactionlab.org/research-resources/introduction-evaluations>
- Karlan, Dean, and Jacob Appel, *Failing in the Field: What We Can Learn When Field Research Goes Wrong*, Princeton, N.J.: Princeton University Press, 2016.
- KPMG International, *Monitoring and Evaluation in the Development Sector, A KPMG International Development Assistance Services (IDAS) Practice Survey*, 2014. As of January 3, 2017:
<https://assets.kpmg.com/content/dam/kpmg/pdf/2014/09/2014-survey-monitoring-evaluation-v4.pdf>
- Kusek, Jody Zall, and Ray C. Rist, *Ten Steps to a Results-Based Monitoring and Evaluation System: A Handbook for Development Practitioners*, Washington, D.C.: World Bank Publications, 2004. As of January 3, 2017:
<https://openknowledge.worldbank.org/bitstream/handle/10986/14926/296720PAPER0100steps.pdf?sequence=1>
- Lamb, Robert D., and Kathryn Mixon, *Rethinking Absorptive Capacity: A New Framework, Applied to Afghanistan's Police Training*, Washington, D.C.: Center for Strategic & International Studies, 2013.
- Liebenthal, Andres, Osvaldo N. Feinstein, and Gregory K. Ingram, “Evaluation and Development: The Partnership Dimension,” *World Bank Series on Evaluation and Development*, Vol. 6, 2004, p. 9.
- Maniruzzaman, Talukder, “Arms Transfers, Military Coups, and Military Rule in Developing States,” *The Journal of Conflict Resolution*, Vol. 36, No. 4, 1992, pp. 733–755.
- Marquis, Jefferson P., Richard E. Darilek, Jasen J. Castillo, Cathryn Quantic Thurston, Anny Wong, Cynthia Huger, Andrea Mejia, Jennifer D. P. Moroney, Brian Nichiporuk, and Brett Steele, *Assessing the Value of U.S. Army International Activities*, Santa Monica, Calif.: RAND Corporation, MG-329-A, 2006. As of August 11, 2017:
<https://www.rand.org/pubs/monographs/MG329.html>
- Marquis, Jefferson P., Michael J. McNerney, S. Rebecca Zimmerman, Merrie Archer, Jeremy Boback, and David Stebbins, *Developing an Assessment, Monitoring, and Evaluation Framework for U.S. Department of Defense Security Cooperation*, Santa Monica, Calif.: RAND Corporation, RR-1611-OSD, 2016. As of August 11, 2017:
https://www.rand.org/pubs/research_reports/RR1611.html
- Marshall, Monty G., and Benjamin R. Cole, “Global Report 2014: Conflict, Governance, and State Fragility,” *Center for Systemic Peace*, 2014. As of January 3, 2017:
<http://www.systemicpeace.org/globalreport.html>
- Mattis, James, “Advance Policy Questions for James N. Mattis, Nominee to Be Secretary of Defense,” 2017. As of August 11, 2017:
https://www.armed-services.senate.gov/imo/media/doc/Mattis%20APQ%20Responses_01-12-17.pdf
- McCormick, James M., and Neil Mitchell, “Is U.S. Aid Really Linked to Human Rights in Latin America?” *American Journal of Political Science*, Vol. 32, No. 1, 1988, pp. 231–239.

McFarlane, Colleen A., "Adjustment Risks of Humanitarian Aid Workers," *The Australian Journal of Disaster and Trauma Studies*, Vol. 2004-1, 2004. As of January 3, 2017:
<http://www.massey.ac.nz/~trauma/issues/2004-1/mcfarlane.htm>

McInnis, Kathleen J., and Nathan J. Lucas, *What Is "Building Partner Capacity"? Issues for Congress*, Washington, D.C.: Congressional Research Service, 2015.

McNerney, Michael J., Jefferson P. Marquis, S. Rebecca Zimmerman, and Ariel Klein, *SMART Security Cooperation Objectives: Improving DoD Planning and Guidance*, Santa Monica, Calif.: RAND Corporation, RR-1430-A, 2016. As of January 3, 2017:
http://www.rand.org/pubs/research_reports/RR1430.html

McNerney, Michael J., Angela O'Mahony, Thomas S. Szayna, Derek Eaton, Caroline Baxter, Colin P. Clarke, Emma Cutrufello, Michael McGee, Heather Peterson, Leslie Adrienne Payne, and Calin Trenkov-Wermuth, *Assessing Security Cooperation as a Preventive Tool*, Santa Monica, Calif.: RAND Corporation, RR-350-A, 2014. As of January 3, 2017:
http://www.rand.org/pubs/research_reports/RR350.html

Meeks, MAJ Donald R., Jr., and MAJ Chad Froehlich, "What Is MPEP?" *Military Police*, Vol. 19-08-2, Fall 2008. As of January 3, 2017:
<http://www.wood.army.mil/engrmag/pdfs/Fall%2008/Meeks-Froehlich.pdf>

Meernik, James, Eric L. Krueger, and Steven C. Poe, "Testing Models of U.S. Foreign Policy: Foreign Aid During and After the Cold War," *The Journal of Politics*, Vol. 60, No. 1, 1998, pp. 63–85.

Mehler, Andreas, "Why Security Forces Do Not Deliver Security: Evidence from Liberia and the Central African Republic," *Armed Forces & Society*, Vol. 38, 2012, pp. 46–69.

Mendelsohn, Joshua, Elizabeth Bodine-Baron, Angela O'Mahony, and Thomas S. Szayna, *A Snapshot of Army Engagement with Partners: Global Landpower Network Baseline*, Santa Monica, Calif.: RAND Corporation, July 28, 2016. Not available to the general public.

Menon, Saraswathi, Judith Karl, and Kanni Wignaraja, *Handbook on Planning, Monitoring and Evaluating for Development Results*, New York: UNDP Evaluation Office, 2009. As of January 3, 2017:
<http://web.undp.org/evaluation/evaluations/handbook/english/documents/pme-handbook.pdf>

Millennium Challenge Corporation, *On the Cutting Edge of Aid Effectiveness. A Primer: Lessons Learned from the Millennium Challenge Corporation*, 2011.

———, "Our Impact," 2016. As of January 3, 2017:
<https://www.mcc.gov/our-impact>

MIT, "Monitoring and Evaluation: What Can They Do for Me?" n.d. As of January 3, 2017:
<http://web.mit.edu/urbanupgrading/upgrading/issues-tools/tools/monitoring-eval.html#Anchor-Monitoring-56567>

Moroney, Jennifer D. P., Beth Grill, Joe Hogler, Lianne Kennedy-Boudali, and Christopher Paul, *How Successful Are U.S. Efforts to Build Capacity in Developing Countries? A Framework to Assess the Global Train and Equipment "1206" Program*, Santa Monica, Calif.: RAND Corporation, TR-1121-OSD, 2011. As of August 11, 2017:
https://www.rand.org/pubs/technical_reports/TR1121.html

Moroney, Jennifer D. P., Joe Hogler, Jefferson P. Marquis, Christopher Paul, John E. Peters, and Beth Grill, *Developing an Assessment Framework for U.S. Air Force Building Partnership Programs*, Santa Monica, Calif.: RAND Corporation, MG-868-AF, 2010. As of August 11, 2017:
<https://www.rand.org/pubs/monographs/MG868.html>

Moroney, Jennifer D. P., Jeff Marquis, Cathryn Quantic Thurston, and Gregory F. Treverton, *A Framework to Assess Programs for Building Partnerships*, Santa Monica, Calif.: RAND Corporation, MG-863-OSD, 2009. As of August 11, 2017:
<https://www.rand.org/pubs/monographs/MG863.html>

Moroney, Jennifer D. P., David E. Thaler, and Joe Hogler, *Review of Security Cooperation Mechanisms Combatant Commands Utilize to Build Partner Capacity*, Santa Monica, Calif.: RAND Corporation, RR-413-OSD, 2013. As of August 11, 2017:
https://www.rand.org/pubs/research_reports/RR413.html

OECD—See Organisation for Economic Co-operation and Development.

- Office of Migrant Education, “Comprehensive Needs Assessment,” New Directors Orientation, 2001. As of January 3, 2017:
<https://www2.ed.gov/admins/lead/account/compneedsassessment.pdf>
- Office of the Secretary of Defense, “Department of Defense Instruction: Assessment, Monitoring, and Evaluation Policy for the Security Cooperation Enterprise,” DoDI 5200.01, 2016.
- O’Mahony, Angela, Thomas S. Szayna, Michael J. McNerney, Derek Eaton, Joel Verneti, Michael Schwille, Stephanie Pezard, Tim Oliver, and Paul S. Steinberg, *Assessing the Value of Regionally Aligned Forces in Army Security Cooperation*, Santa Monica, Calif.: RAND Corporation, RR-1341/1-A, 2017a.
- O’Mahony, Angela, Thomas S. Szayna, Christopher G. Pernin, Laurinda L. Rohn, Derek Eaton, Elizabeth Bodine-Baron, Joshua Mendelsohn, Osonde A. Osoba, Sherry Oehler, Katharina Ley Best, and Leila Bighash, *Global Landpower Network: Recommendations for Strengthening Army Engagement*, Santa Monica, Calif.: RAND Corporation, RR-1813-A, 2017b.
- Operational Policy and Quality Department, *Results Framework and M&E Guidance Note*, Washington, D.C.: World Bank, April 9, 2013. As of January 3, 2017:
http://siteresources.worldbank.org/PROJECTS/Resources/40940-1365611011935/Guidance_Note_Results_and_M&E.pdf
- Organisation for Economic Co-operation and Development, *The Paris Declaration on Aid Effectiveness and the Accra Agenda for Action*, 2005/2008. As of January 3, 2017:
<http://www.oecd.org/dac/effectiveness/34428351.pdf>
- Organisation for Economic Co-operation and Development-Development Assistance Committee, *Principles for Evaluation of Development Assistance*, Paris: OECD, 1991. As of January 3, 2017:
<https://www.oecd.org/development/evaluation/dcdndep/41029845.pdf>
- Parsons, Jim, Caitlin Gokey, and Monica Thornton, *Indicators of Inputs, Activities, Outputs, Outcomes and Impacts in Security and Justice Programming*, UK Aid, Department for International Development, October 15, 2013. As of January 3, 2017:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/304626/Indicators.pdf
- Partnerships Group—Strategy and Resource Management, “Partnerships for Development: Proposed Actions for the World Bank, a Discussion Paper,” *The World Bank*, May 20, 1998. As of January 3, 2017:
<http://documents.worldbank.org/curated/en/789141468153858137/text/516050WP0Dev0i10Box342046B01PUBLIC1.txt>
- Paul, Christopher, “Foundations and Assessment: The Hierarchy of Evaluation and the Importance of Articulating a Theory of Change,” *Small Wars Journal*, July 30, 2013.
- Paul, Christopher, Colin P. Clarke, Beth Grill, Stephanie Young, Jennifer D. P. Moroney, Joe Hogler, and Christine Leah, *What Works Best When Building Partner Capacity and Under What Circumstances?* Santa Monica, Calif.: RAND Corporation, MG-1253/1-OSD, 2013a. As of January 3, 2017:
<http://www.rand.org/pubs/monographs/MG1253z1.html>
- Paul, Christopher, Michael Nixon, Heather Peterson, Beth Grill, and Jessica Yeats, *The RAND Security Cooperation Prioritization and Propensity Matching Tool*, Santa Monica, Calif.: RAND Corporation, TL-112-OSD, 2013b. As of August 11, 2017:
<https://www.rand.org/pubs/tools/TL112.html>
- Ploch, Lauren, “Africa Command: U.S. Strategic Interests and the Role of the U.S. Military in Africa,” *Congressional Research Services*, July 2011. As of January 3, 2017:
<https://www.fas.org/srg/crs/natsec/RL34003.pdf>
- Rajan, Raghuram G., and Arvind Subramanian, “Aid and Growth: What Does the Cross-Country Evidence Really Show?” *The Review of Economics and Statistics*, Vol. 90, No. 4, 2008, pp. 643–665.
- Rand, Dafna H., and Stephen Tankel, *Security Cooperation & Assistance: Rethinking the Return on Investment*, Washington, D.C.: Center for a New American Security, 2015.
- Rao, Vijayendra, and Michael Woolcock, “Integrating Qualitative and Quantitative Approaches in Program Evaluation,” in *The Impact of Economic Policies on Poverty and Income Distribution: Evaluation Techniques and Tools*, Washington, D.C.: World Bank and Oxford University Press, 2003, pp. 165–190. As of January 3, 2017:
http://siteresources.worldbank.org/INTPSIA/Resources/490023-1121114603600/12930_Chapter8.pdf

Rawlings, Laura, "Who's Afraid of Administrative Data? Why Administrative Data Can Be Faster, Cheaper and Sometimes Better," *Development Impact Blog*, World Bank, June 26, 2013. As of January 3, 2017: <http://blogs.worldbank.org/impacitevaluations/whos-afraid-administrative-data-why-administrative-data-can-be-faster-cheaper-and-sometimes-better>

Roberts, Dawn, and Nidhi Kattri, *Designing a Results Framework for Achieving Results: A How-To Guide*, Independent Evaluation Group, Washington, D.C.: World Bank, 2012.

Ruby, Tomislav Z., and Douglas Gibley, "US Professional Military Education and Democratization Abroad," *European Journal of International Relations*, Vol. 16, No. 3, 2010, pp. 339–364.

Shadish, William R., Thomas D. Cook, and Donald T. Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Boston: Cengage Learning, 2002.

Shah, Neil Buddy, Paul Wang, Andrew Fraker, and Daniel Gastfriend, "Evaluations with Impact: Decision-Focused Impact Evaluation as a Practical Policymaking Tool," 3ie International Initiative for Impact Evaluation, Working Paper #25, September 2015. As of January 3, 2017: <http://www.3ieimpact.org/en/publications/working-papers/3ie-working-paper-25/>

Sislin, John, "Arms as Influence: The Determinants of Successful Influence," *The Journal of Conflict Resolution*, Vol. 38, No. 4, 1994, pp. 665–689.

"Southern Accord 16," *U.S. Army Africa*, August 1, 2016. As of January 3, 2017: <https://www.army.mil/standto/2016-08-01>

Stern, Elliot D., Laura Altinger, Osvaldo Feinstein, Marta Marnon, Daniela Ruegenberg, Nils-Sjard Schultz, and Nicolai Steen Nielsen, "Thematic Study on the Paris Declaration, Aid Effectiveness and Development Effectiveness," *Evaluation of the Paris Declaration*, November 2008. As of January 3, 2017: <https://www.oecd.org/dac/evaluation/dcdndep/41807824.pdf>

TRADOC, *U.S. Army Functional Concept for Engagement*, TRADOC Pamphlet 525-8-5, February 2014a.

———, *The U.S. Army Operating Concept: Win in a Complex World, 2020-2040*, TRADOC Pamphlet 525-3-1, October 2014b.

United Nations Development Programme, "Innovations in Monitoring & Evaluating Results, Discussion Paper," November 5, 2013. As of January 3, 2017: [http://www.undp.org/content/dam/undp/library/capacity-development/English/Discussion%20Paper-%20Innovations%20in%20Monitoring%20&%20Evaluating%20Results%20\(5\).pdf](http://www.undp.org/content/dam/undp/library/capacity-development/English/Discussion%20Paper-%20Innovations%20in%20Monitoring%20&%20Evaluating%20Results%20(5).pdf)

United States Africa Command, "Flintlock," n.d. As of January 3, 2017: <http://www.africom.mil/what-we-do/exercises/flintlock>

U.S. Agency for International Development, *U.S. Overseas Loans and Grants: Obligations and Loan Authorizations, July 1, 1945–September 30, 2014*, n.d. As of January 3, 2017: <https://explorer.usaid.gov/reports-greenbook.html>

U.S. Congress, *National Defense Authorization Act for Fiscal Year 2016*, n.d.a. As of January 3, 2017: <https://www.congress.gov/bill/114th-congress/house-bill/1735>

———, *National Defense Authorization Act for Fiscal Year 2017*, n.d.b. As of January 3, 2017: <https://www.congress.gov/114/crpt/hrpt537/CRPT-114hrpt537.pdf>

U.S. Department of State, *Voting Practices in the United Nations 2014: Report to Congress Submitted Pursuant to Public Laws 101-246 and 108-447*, July 2015. As of January 3, 2017: <http://www.state.gov/documents/organization/245163.pdf>

USAID, *Evaluation, Learning from Experience: USAID Evaluation Policy*, January 2011. As of January 3, 2017: <https://www.usaid.gov/sites/default/files/documents/2151/USAIDEvaluationPolicy.pdf>

———, *Strengthening Evidence-Based Development, Five Years of Better Evaluation Practice at USAID 2011-2016*, 2016a. As of January 3, 2017: <https://www.usaid.gov/sites/default/files/documents/1870/Strengthening%20Evidence-Based%20Development%20-%20Five%20Years%20of%20Better%20Evaluation%20Practice%20at%20USAID.pdf>

———, "USAID Evaluation," 2016b. As of January 3, 2017: <https://www.usaid.gov/evaluation>

- Valters, Craig, "Theories of Change: Time for a Radical Approach to Learning in Development," Overseas Development Institute, September 2015. As of January 3, 2017:
<https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9835.pdf>
- Vogel, Isabel, *Review of the Use of 'Theory of Change' in International Development*, Review Report, United Kingdom: Department for International Development (DFID), 2012. As of January 3, 2017:
https://assets.publishing.service.gov.uk/media/57a08a5ded915d3cfd00071a/DFID_ToC_Review_VogelV7.pdf
- Wang, Miao, and M. C. Sunny Wong, "FDI, Education, and Economic Growth: Quality Matters!" *Atlantic Economic Journal*, Vol. 39, No. 2, 2011, pp. 103–115.
- Watkins, Ryan, and Maurya West Meiers, *A Guide to Assessing Needs: Essential Tools for Collecting Information, Making Decisions, and Achieving Development Results*, New York: World Bank Publications, 2012. As of January 3, 2017:
<https://openknowledge.worldbank.org/handle/10986/2231>
- Watts, Stephen, Christopher M. Schnaubelt, Sean Mann, Angela O'Mahony, and Michael Schwille, *Pacific Engagement: Forging Tighter Connections Between Tactical Security Cooperation Activities and U.S. Strategic Goals in the Asia-Pacific Region*, Santa Monica, Calif.: RAND Corporation, RR-1920-A, 2017.
- Western Hemisphere Institute for Security Cooperation, *Course Catalog 2017–2018*, March 17, 2015. As of January 3, 2017:
[http://www.benning.army.mil/tenant/whinsec/content/pdf/WHINSEC%202017-2018%20Course%20Catalog%20\(English\).pdf](http://www.benning.army.mil/tenant/whinsec/content/pdf/WHINSEC%202017-2018%20Course%20Catalog%20(English).pdf)
- Whitaker, Beth Elise, "Compliance Among Weak States: Africa and the Counter-Terrorism Regime," *Review of International Studies*, Vol. 36, 2010, pp. 639–662.
- The White House, *National Security Strategy*, February 2015. As of January 3, 2017:
https://www.whitehouse.gov/sites/default/files/docs/2015_national_security_strategy.pdf
- Wikipedia, "List of Countries by English-Speaking Population," n.d. As of January 17, 2017:
https://en.wikipedia.org/wiki/List_of_countries_by_English-speaking_population
- Williams, Lt. Col. Antwan C., "U.S. Army South Conducts Army-to-Army Staff Talks with Chile," *U.S. Army South Public Affairs*, October 31, 2011. As of January 3, 2017:
https://www.army.mil/article/68328/U_S__Army_South_conducts_army_to_army_staff_talks_with_Chile
- W. K. Kellogg Foundation, *W. K. Kellogg Foundation Evaluation Handbook*, January 1, 2010. As of January 3, 2017:
<https://www.wkkf.org/resource-directory/resource/2010/w-k-kellogg-foundation-evaluation-handbook>
- World Bank, *Monitoring and Evaluation: Some Tools, Methods & Approaches*, Operations Evaluation Department, Washington, D.C.: World Bank, 2004.
- , "Country Policy and Institutional Assessment. Frequently Asked Questions," 2011. As of January 3, 2017:
<http://go.worldbank.org/EEAIU81ZG0>
- , *World DataBank*, 2016. As of January 3, 2017:
<http://databank.worldbank.org/data/home.aspx>
- Young, Cristobal, "Model Uncertainty in Sociological Research: An Application to Religion and Economic Growth," *American Sociological Review*, Vol. 74, No. 3, 2009, pp. 380–397.

