

Lessons Learned from Stabilization Initiatives in Afghanistan: A Systematic Review of Existing Research

Radha Iyengar¹ Jacob N. Shapiro² Stephen Hegarty³

April 11, 2017

This research was conducted by the Empirical Studies of Conflict Project (ESOC) at Princeton University. The authors and ESOC are grateful for the support provided by the United States Agency for International Development (USAID) and the United States Institute of Peace (USIP). USIP convened and chaired several Advisory Board meetings on this project, and the inputs and guidance provided by independent Advisory Board members is gratefully acknowledged. USAID helped organize interviews and provided unprecedented access to internal data. Funding was provided through an interagency research agreement from the Office of Afghanistan and Pakistan Affairs at USAID and USIP. All errors are our own.

¹ Senior Economist, RAND Corporation; Visiting Research Scholar, Princeton University

² Professor of Politics and International Affairs, Princeton University, Co-Director, Empirical Studies of Conflict Project (ESOC)

³ Research Specialist, Princeton University

Contents

Acronyms	4
Executive Summary	6
1 INTRODUCTION	11
2 APPROACH	12
2.1 Study Selection	12
2.2 Defining Stabilization	14
2.3 Assessment Criteria	15
3 KEY FINDINGS	18
3.1 Stabilization Programs Impact on Key Outcomes	19
3.1.1 Security	19
3.1.2 Support for Government and Anti-Government Elements	20
3.1.3 Community Cohesion and Resilience	20
3.1.4 Health and Economic Well-Being	21
3.2 Time Horizon of Effects	21
3.2.1 Limited Evidence of Short-Term Impact	22
3.2.2 Gaps in Understanding Long-Term Impact	23
3.3 Impact of Military Presence	23
3.3.1 International Military Forces have a Role in the Basic Execution of Stabilization Programs	24
3.3.2 Civilian and Military Implementers Have Conflicting Objectives	24
3.3.3 The Securitization of Development Aid is a Controversial Concept	24
3.4 Synergies and Confounding Factors with Other Donors	25
3.5 Impact of Specific Project Aspects	26
3.5.1 Small-Scale Programs Produced Quick Impact Results, but Did Not Address Drivers of Instability	26
3.5.2 Large-Scale Programs Created Unrealistic Expectations and Experienced Higher Levels of Corruption and Targeting	27

3.6	Summary of Commonalities across Successful or Unsuccessful Projects	27
3.6.1	The Afghan National Government’s Commitment to Reform is a Fundamental Prerequisite for Success	28
3.6.2	Impact of Corruption is Pervasive and Corrosive	28
3.6.3	Security is a Key Determinant of Program Success and Sustainability	28
3.6.4	Employing Competent, Long-term, and, Ideally, Local Staff are Essential Elements of Successful Projects	29
4.	CONCLUSIONS AND RECOMMENDATIONS	30
	REFERENCES	31

Acronyms

ACSOR	Afghan Center for Socio-Economic and Opinion Research
AGE	Anti-Government Elements
AIMS	Aid Information Management Systems
AISCS	Afghanistan Infrastructure and Security Cartography System
ALLI	Alternative Licit Livelihoods Initiatives
ANDP	Afghanistan National Development Program
ANQAR	Afghanistan Nationwide Quarterly Assessment Research
ANSO	Afghanistan NGO Safety Office
ANVIL	Name of survey (not an acronym)
ASI	Afghanistan Stability Initiative
BINNA	Name of survey (not an acronym)
CBSG	Community Based Stabilization Grants
CCI	Community Cohesion Initiative Project
CDP	Community Development Program
CERP	Commander's Emergency Response Program
CIDA	Canadian International Development Agency
CSO	Central Statistics Organization
DFID	Department for International Development
DHS	Demographic and Health Surveys
DOD	Department of Defense
DTEM	Digital Terrain Elevation Map
ESOC	Empirical Studies of Conflict Project
FOB/COP	Forward Operating Base/Combat Outpost
FOGHORN	Name of survey (not an acronym)
GIRoA	Government of the Islamic Republic of Afghanistan
GIS	Geographic Information Systems
ISAF	International Security Assistance Force
ISVG	Institute for the Study of Violent Groups
LGCD	Local Governance and Community Development Project
MICS	Multiple Indicator Cluster Survey
MISTI	Measuring the Impact of Stabilization Initiatives Project
MRRD	Ministry for Rural Rehabilitation and Development
NATO	North Atlantic Treaty Organization

NGO	Non-Governmental Organization
NOAA	National Oceanic and Atmospheric Administration
NRVA	National Risk and Vulnerability Assessment
NSP	National Solidarity Program
NTMA	NATO Training Mission in Afghanistan
OAPA	USAID Office of Afghanistan and Pakistan Affairs
OTI	USAID Office of Transition Initiatives
PAP	Pre-Analysis Plan
SIGACTS	Significant Activities (e.g., violent events)
SIKA	Stability in Key Areas Project
STAY	Skills Training for Afghan Youth
UNDSS	United Nations Department for Safety and Security
USAID	United States Agency for International Development
USG	United States Government
USIP	United States Institute for Peace

Executive Summary

This report summarizes findings from a review of 89 studies on development and stabilization programming in Afghanistan. These findings inform answers to the following six research questions that were identified before conducting the research review:⁴

R1: What did stabilization projects achieve in terms of key outcomes, including: security; popular support for the government; popular support for anti-government elements (AGE); community cohesion and resilience; health of the Afghan people; economic well-being of the Afghan people; and conflict events?

R2: Over what time horizon is these effects apparent and how quickly do any gains or losses fade?

R3: How does the presence of the military impact the outcomes of stabilization projects?

R4: What types of synergies and confounding factors exist between stabilization programs by different actors (other parts of the United States Government (USG), other countries, the Afghan government, international organizations like the World Bank, etc.)?

R5: Are impacts of stabilization programs amplified or reduced when considering specific aspects (size, contract type, etc.) or sectors (agriculture, infrastructure, skills, etc.) of projects?

R6: What commonalities exist when looking across a number of successful or unsuccessful stabilization projects between different actors and different sectors?

This report summarizes the high-level findings that cross-cut these six questions as well as the specific evidence related to each of the questions. For each of the 89 studies, we applied three criteria identified in a pre-analysis plan, which helped make the analysis systematic and reduced bias: internal validity of causal claims; scope of coverage; and stabilization indicators. We established these assessment criteria before beginning to review these studies in order to focus the analysis on the most relevant issues and findings using unbiased standards to compare sources. The motivations of our study are to identify the key factors that contributed to the success or failure of stabilization programs in Afghanistan and formulate recommendations for the design, implementation and evaluation of future efforts in Afghanistan and other conflict-

⁴ These questions were modified slightly during the course of the project to address feasibility constraints from the existing data and literature. These changes were detailed in an addendum to the Pre-Analysis Plan (PAP).

affected areas.

Across the literature, we find some evidence of near-term gains in security, health access, and economic activity (R1). The small, short-term improvement in local security varied considerably across region and program and there are notable examples where security did not improve and even worsened. In all cases, the estimated effect was small, especially relative to the overall rates of violence. This variation is a common theme across a number of the outcomes. Support for the Afghan government and AGE varies substantially across different programs making it difficult to draw a unified conclusion across the literature. Much of the variation in those attitudes appears driven by perceptions of government corruption. There was no evidence to suggest universal belief that insurgents were better at governance or that support was zero-sum with Afghan Government. The evidence on community cohesion in the existing literature was too limited to draw a conclusion and in many studies was not even considered. There was some evidence of improved health services (both actual and perceived access) but few measures of actual health improvement to indicate whether this improved access translates into improved well-being.⁵ There was also evidence of consistent economic gains. Some of this may have been short-lived due to increased spending and activity during the project.

The evidence that any of these projects had a long-term impact is even more limited (R2). This is in part because few studies were designed with prospective or retrospective design to estimate sustainability. Some projects had unrealistic goals and mechanisms to facilitate sustainable outcomes (e.g., depended on continued spending or international support) and others suffered from institutional limitations (e.g., corruption, lack of capacity) and were unlikely to see sustained gains. There is also some evidence that military forces played a key role in facilitating basic operations (R3). However, military and civilian implementers had different timelines and objectives making coordination challenging. There was also a core set of scholars who worried that the presence of the military as an implementer, or even simply providing security, “militarized” aid provision; thus inhibiting it from achieving social welfare objectives. In contrast, there were few evaluations on how to achieve successful coordination between international donors. However, multiple studies noted the importance of coordination between different countries and the host nation in ensuring success (R4).

Overall, the most important program feature that could enable success was program size (R5). The literature consistently found that small-scale programs produced near-term impact on at least some of the key indicators. However, because of their limited size and scope, these programs could not address key drivers of instability and thus were often not associated with longer-term

⁵ This includes a broad range of health care services measures from the literature. We note that this differs from the more limited measure of perceived access to health care discussed in detail in the Iyengar, Shapiro, Mao, and Singh (2017) paper.

changes. However, large-scale programs created unrealistic expectations and were more subject to corruption and targeting by insurgents, undermining their near-term effectiveness. The literature is also very informative on key enabling factors (R6). Key factors that enabled success include: host-nation coordination and commitment, limiting the extent of corrosive corruption, ensuring baseline levels of security to facilitate basic implementation and oversight, and ensuring appropriate staffing, both in terms of skills and in terms of longevity of deployment were all associated with programmatic success.

Looking across the 89 studies, there were four key themes that cross-cut the six analytic questions. First, most stabilization programs will have – at best – modest impact (less than 0.1 standard deviation when measured quantitatively and typically described as small in qualitative analysis). Based on the Afghanistan experience, policy makers and implementers should not expect to generate either large or persistent effects. From well-designed experimental and quasi-experimental approaches (e.g., Beath, Fontini and Enikolopov, 2013; MSI, 2014) to government-initiated qualitative reviews (Bohnke, Koehler and Zurcher, 2014; Bohnke and Zurcher, 2013 (1b); Norad, 2012) to historical accounts (Goodhand, 2002) the evidence consistently indicates stabilization programming has small, generally transitory, impacts (both positive or negative). Programs such as the Commander’s Emergency Response Program (CERP) (Chou, 2012; Sexton, 2015) or some of the Afghanistan Stabilization Initiative (ASI) programs (Altai, 2012 (1)) that have been “successful” may have short-term positive impacts, but they do not appear to generate large shifts in security, attitudes, or capacity. This is relevant for both managing expectations for what stabilization programs may accomplish and for considering how to design measurement and evaluation efforts to detect relatively small effects.

Second, smaller may be better. A number of studies (Altai, 2012 (7); Chou, 2012; Child, 2014; Goodhand, 2002; Gordon, 2011; Kapstein and Kathuria, 2012; Nagl, Exum and Humayun, 2009) highlight the intuition that smaller projects can be targeted at important, specific gaps and seem less likely to fuel instability. Small projects have a variety of beneficial features: they are often easier to manage by staff on the ground; they are less likely than large infrastructure projects to attract attention from corrupt officials or to become targets for enemy sabotage; and outputs are small and less likely to become a source of conflict. The literature does not provide evidence of increasing returns from a cost-effectiveness perspective: small projects do not appear to have a differential impact on outcomes such as violence or support for the government relative to larger-scale projects (see Child, 2014). Additionally, Afghans reacted positively to large-scale programs that are populated by small, community driven projects such as the Afghanistan National Solidarity Program (NSP) at least in part because the funding dispersed was too small to be siphoned off by powerful interests, though it still did provide meaningful benefits to communities (Gordon, 2011). This is also important for minimizing the risk of unintended negative outcomes. For instance, any corruption that affects community-driven, local projects is

by its nature smaller in scale and thus less likely to delegitimize the national government (Kapstein and Kathuria, 2012).

Third, stabilization efforts should be designed in ways that make it hard for destabilizing forces to target or claim credit for programs (Altai, 2012 (1-3); Carbonnier, 2014; IMU, 2015 (2)). This insight is particularly important for interpreting the variation in outcomes when small projects have both small, positive effects (Beath, Fontini and Enikolopov, 2013 (1a)) and some small, negative impacts (e.g., MSI, 2014). There are two important mechanisms that can generate negative impacts from otherwise potentially effective programs: one, programs can be deliberately targeted and de-legitimized by insurgents (e.g., Altai, 2012 (5); Lyall, 2016; Sexton, 2015); or, two, AGE may take credit for positive effects or seek bribes and revenue from the programs; this tends to raise perceptions of corruption, for which Afghans almost automatically blame the national government, thereby reducing its legitimacy and increasing support for AGE (e.g., MSI, 2014; ICG, 2011). Programs must therefore not only be able to effectively address short-term drivers of instability, but also must be properly credited to the government or local leaders without attracting violence or other negative attention from AGE. One effective strategy, successfully employed by the ASI program, is to ensure local government and local non-governmental organizations (NGOs) are visibly central to project implementation (Altai, 2012(7)).

Fourth, and finally, while we have systematically compiled the evidence from the existing literature, almost none of the ideas presented in this review are new. Many of the studies included in this analysis noted findings based on the evidence available several years ago – but few of those recommendations have been implemented. Many of our findings regarding development impacts based on studies through 2016 are also highlighted in ICG (2011), which argued: “The impact of international assistance will remain limited unless donors, particularly the largest, the U.S., stop subordinating programming to counter-insurgency objectives, devise better mechanisms to monitor implementation, adequately address corruption and wastage of aid funds, and ensure that recipient communities identify needs and shape assistance policies.” Of particular note is the repeated, widespread recommendation for improved monitoring and evaluation, which is succinctly summarized by Bohnke, Koehler and Zurcher (2014) as follows: “...if the international community is serious about rigorous impact evaluations, it must pressure donors and implementing actors for much higher standards for recording and sharing data!” Many of these recommendations could be achieved through adopting the improvements and changes suggested in Department of State (2011).

We must note here that it is unsurprising that many programs did not accomplish the desired outcomes; few were designed, implemented or modified to take into account existing recommendations that might have improved their chances for success. We acknowledge the

pervasive impact of environmental factors in Afghanistan (e.g., security, corruption) on program success. However, it is precisely because stabilization efforts in complex and difficult environments are inherently dynamic processes that future efforts should focus on not simply implementing projects but on ensuring a mechanism for effectively integrating evidence-based recommendations and, when appropriate, modifying policy and strategy to account for empirical findings.

1 INTRODUCTION

While there are multiple studies on the effects of specific stabilization programs in Afghanistan, there is no study that systematically reviews the full range of studies on stabilization and development programming in the country. Such a review is useful because while any individual study may have specific limitations, common themes identified across the range of studies are likely to be broadly accurate. In particular, since each study uses a slightly different methodology and is done by different organizations with varying views and methods, looking across studies effectively cancels out biases and should reveal more reliable patterns. This is the core logic behind meta-analysis in the sciences and it applies in this setting as well.

This document summarizes key findings from program evaluations, government documents, the academic literature, and policy studies, on the efficacy and impact of stabilization efforts in Afghanistan. Specifically, this review compiles and analyzes evidence on the following six research questions:

R1: What did stabilization projects achieve in terms of key outcomes: security; popular support for the government; popular support for anti-government elements (AGE); community cohesion and resilience; health of the Afghan people; economic well-being of the Afghan people; and conflict events?

R2: Over what time horizon is these effects apparent and how quickly do any gains or losses fade?

R3: How does the presence of the military impact the outcomes of stabilization projects?

R4: What types of synergies and confounding factors exist between stabilization programs by different actors (other parts of the USG, other countries, Afghan government, international organizations like the World Bank, etc.)?

R5: Are impacts of stabilization programs amplified or reduced when considering specific aspects (size, contract type, etc.) or sectors (agriculture, infrastructure, skills, etc.) of projects?

R6: What commonalities exist when looking across a number of successful or unsuccessful stabilization projects between different actors and different sectors?

The rest of this document summarizes the methodological approach and key findings for each of

these six questions. The final section summarizes key cross-cutting findings relevant for policy makers and provides some recommendations for future policy and research regarding stabilization programs in conflict-affected areas.

2 APPROACH

2.1 Study Selection

The literature reviewed for the study was identified through the research team’s search of the academic and think tank literature on the impact of aid programs in Afghanistan, which was supplemented by government reports recommended by USAID and documents cited and referenced in the research reviewed. The research team prioritized studies with a specific focus on USAID stabilization efforts in Afghanistan in the last decade; a secondary priority was on literature assessing Afghanistan stabilization programming implemented by other USG agencies and international donors. Comparative studies of stabilization initiatives in countries other than Afghanistan were included if they addressed the main relevant themes of the research review.

In total, the research team examined 110 studies and selected a total of 89 for inclusion in the review. Of these studies, 36 focus explicitly on analyzing USAID or Office of Transition Initiatives (OTI) administered programs. Another 11 studies focus on other USG activities, such as US military stabilization programs under the Commander’s Emergency Reconstruction Program (CERP) and operations of Provincial Reconstruction Teams (PRTs). We then identified about 25 additional documents based on citations and references in the research reviewed. The team excluded several studies that were theoretical or “think pieces” and were not intended to serve an evaluation function. We also excluded pieces that only tangentially addressed Afghanistan (e.g., those with a focus on Iraq or the Philippines), but retained reports that included Afghanistan as a substantive case even if other countries were included in the analysis.⁶

To structure our review, we organized the literature into four broad categories⁷:

1. **Program Evaluation.** Comprehensive review of USAID-commissioned studies of the impact and effectiveness of its stabilization programs in Afghanistan (to include specialized thematic evaluations of these programs, specific to gender, ethnicity, region etc.). Examples include:

⁶ The findings and analysis are based on the review presented in Appendix A, which provides a detailed summary of each study included. Each study is assigned a unique ID for cross-reference between the Summary Report and Reference Table. This report uses these unique IDs to reference any findings and conclusions.

⁷ These categories were included in June 2016 pre-analysis plan provided to and approved by USAID OAPA.

- a. Social Impact. February, 2016. *Final Performance Evaluation of the USAID/OTI Community Cohesion Initiative*.
 - b. Management Systems International. February, 2015. *Afghan Civilian Assistance Program II: Final Performance Evaluation*.
 - c. The Measuring the Impact of Stabilization Initiatives Project (MISTI) Stabilization Trends and Impact Evaluation Survey.
2. **Government Documents.** Official USG and foreign government reviews or assessments of the efficacy of stabilization assistance programs in Afghanistan. Examples include:
 - a. Special Inspector General for Afghanistan Reconstruction (SIGAR). March, 2011. *Afghanistan's National Solidarity Program Has Reached Thousands of Afghan Communities, but Faces Challenges that Could Limit Outcomes*.
 - b. UKAID Stabilisation Unit. November, 2010. *Stabilisation Case Study: Infrastructure in Helmand, Afghanistan*.
3. **Academic Literature.** Review of social science scholarship (books and peer-reviewed journal articles) on aid and stabilization in Afghanistan published in the last decade. Examples include:
 - a. Gordon, Stuart. 2014. "Afghanistan's Stabilization Program: Hope in a Dystopian Sea?" in Robert Muggah, ed. *Stabilization Operations, Security and Development: States of Fragility*. New York: Routledge.
 - b. Sexton, Renard. 2015. "Aid as a Tool against Insurgency: Evidence from Contested and Controlled Territory in Afghanistan." *American Political Science Review*.
4. **Policy/Think Tank.** Independent studies of aid and stabilization in Afghanistan with a policy-centered orientation. Examples include:
 - a. Viehe, Ariella, Jasmine Afshar and Tamana Heela. December, 2015. *Rethinking the Civilian Surge: Lessons from the Provincial Reconstruction Teams in Afghanistan*. Washington, DC: Center for American Progress.
 - b. Kapstein, Ethan and Kamna Kathuria. December, 2012. *Economic Assistance in Conflict Zones: Lessons from Afghanistan*. Washington, DC: Center for Global Development.

These categories served two key functions. First, they enabled the research team to compare studies that were similar in purpose and design. For instance, we could compare program evaluations to other evaluations rather than to more qualitative think tank reports. This comparison allowed us to assess and evaluate both methods and findings in the broader context of the purpose and audience of the report. Second, this allowed us to collate and integrate information within and across study types to provide a more holistic review of the findings. These categories are not intended to provide any ranking or normative weight on the relative value of any specific type of study. Our subsequent analysis found studies that were classified as

carefully designed and effective evaluations in each of the four categories of studies.

2.2 Defining Stabilization

The study explicitly posits no normative definition of stabilization. This is because there is a substantial degree of variation in this definition and we did not want to inadvertently exclude certain research or analysis through the initial definitional choice. Thus, the research team instead allowed the definition to emerge from an analysis of stabilization indicators in policy documents, program evaluations and academic literature. Some of the most common types of aid labelled “stabilization programming” that we encountered in our review include:

- Efforts to improve local government capacity for service delivery to increase legitimacy and strengthen ties with local communities
- Community-led small infrastructure projects to improve community cohesion and resilience to conflict
- Youth training and education to increase positive engagement with the community and reduce susceptibility to violent extremism
- Agricultural development to provide alternatives to poppy cultivation
- Short-term employment generation efforts often called “cash for work” programs.

Across these studies, we found nearly 200 different indicators used in various combinations to measure and track implicit or explicit definitions of stabilization. Based on these indicators, we developed a classification system of eight broad factors associated with stabilization. These were the key factors and associated indicators that occurred regularly throughout the literature and analyzing along these dimensions allowed us to compare findings between studies with common themes but different specific measures or metrics. The factors are: attitudes towards the Afghan government (including government at any level and civil society attitudes broadly such as support for voting, government-run institutions, and views on national identity); attitudes towards anti-government elements (AGE), including criminal and insurgent groups; attitudes towards foreign actors (e.g., military forces, development implementers); economic well-being; government capacity; health and social well-being; infrastructure improvement; and social cohesion. The most common indicators were security, government capacity, and attitudes towards the Afghan government. Less common indicators were infrastructure improvements and attitudes towards AGE. Rarely were either economic well-being or health and social well-being included as indicators.

Table 1. Number of Studies with Key Indicators

Categories	Count of Studies
Government Capacity	57
Attitudes Towards Afghan Government	41
Security	41
Social Cohesion	26
Attitudes Towards AGE	11
Economic Well-Being	10
Attitude Towards Foreign Actors	9
Infrastructure Improvements	9
Health and Social Well-Being	5

A compilation of stabilization indicators included in the Reference Table (attached as an appendix) highlights an underlying theme of this review: in the complex and fraught environment aid implementers encountered in Afghanistan, defining stabilization - much like designing and implementing stabilization programs - was a variable, sometimes vague, and highly dynamic process. As such, the research team coded the nature of indicators used in each of the studies and then assessed studies across these measures without judgement as to whether one definition of stabilization (and associated combination of indicators) was preferable to another.

2.3 Assessment Criteria

Every study included for analysis was assigned a classification for three assessment criteria: the internal validity of the causal claims made; the scope of coverage (both temporal and geographic) offered; and the framework of stabilization indicators employed.⁸ The categories for these criteria can be found in Table 2, below.

⁸ These categories were included in June 2016 pre-analysis plan provided to and approved by USAID OAPA.

Table 2. Research Review Assessment Criteria

Evaluation Criteria	Effective	Potentially Effective	Insufficient Evidence
Internal Validity of Causal Claims	Experimental: random selection of treatment and control groups; robust estimates of causal effects; defensible and scope for generalizability clearly discernible	Systematic: purposive or random sampling or observations; data collection is systematic; plausible attribution of causal effects based on research design; thorough exploration of correlations	Anecdotal: no strong methodology; data collection is opportunistic; unreliable or idiosyncratic impressions; unable to establish causal effects
Scope of Coverage	Complete geographic coverage of program area; multiple waves of data collection; baseline data employed and outcomes in control areas measured	Complete or nearly complete geographic coverage of program area; examines a single time period and/or is quantitative or qualitative only (not mixed method); no baseline data	Partial geographic coverage of program area; examines a single, relatively short time period
Stabilization Indicators	Rigorous framework of stability indicators	Plausible definition of stability indicators	Missing or vague definition of stability indicators

It is important to note that these assessment criteria do not exclude or negatively code qualitative studies. Qualitative studies with well-defined research questions, defined indicators or measures, and a careful interview or document review design were coded as high quality. Ultimately, our review includes both quantitative studies that lack the rigorous and systematic approach needed to provide credible estimates, as well as well-designed qualitative studies that provide credible evidence on the effects of programs and critical insights that could not be collected through quantitative measures.

We utilize these criteria to enable a more nuanced assessment of each study’s utility. For example, a study found to be “Effective” in terms of analytical rigor, but with “Insufficient

Evidence” related to its scope of coverage may provide a highly credible, but potentially highly localized assessment of stabilization programming impact, with findings that are not readily generalizable to broader environments or contexts.

In addition to the pre-specified criteria, we developed a coding for “classification of practical findings or recommendations” to highlight some of the studies’ more practical suggestions on the design and implementation of stabilization programs. This category has three broad dimensions: information can be characterized as having limited value for current planning or insufficient practical application; some useful findings or recommendations; or useful findings or recommendations. For each of these categories, we separated the findings into four areas: program design, program implementation and oversight, measurement and evaluation, and civil-military coordination.

3 KEY FINDINGS

Many of the 89 studies reviewed relied wholly or in part on quantitative data. Many studies also included interviews with government officials in Afghanistan, civilian personnel from ISAF nations, military personnel from ISAF nations, and Afghan civilians. The studies were largely conducted independently from each other, often without reference to each other in different fields and for different audiences.

The rest of this chapter details the findings from the reports on six specific research questions:

R1: What did stabilization projects achieve in terms of key outcomes: security; popular support for the government; popular support for anti-government elements (AGE); community cohesion and resilience; health of the Afghan people; economic well-being of the Afghan people; and conflict events?

R2: Over what time horizon is these effects apparent and how quickly do any gains or losses fade?

R3: How does the presence of the military impact the outcomes of stabilization projects?

R4: What types of synergies and confounding factors exist between stabilization programs by different actors (other parts of the USG, other countries, Afghan government, international organizations like the World Bank, etc.)?

R5: Are impacts of stabilization programs amplified or reduced when considering specific aspects (size, contract type, etc.) or sectors (agriculture, infrastructure, skills, etc.) of projects?

R6: What commonalities exist when looking across a number of successful or unsuccessful stabilization projects between different actors and different sectors?

The evidence presented in each section highlights areas where there is broad consensus based on effective evidence (“current knowns”); widely held ideas that are not supported by effective evidence (“current assumptions”); research questions or topics raised in the literature for which empirical evidence does not exist (“current gaps”); and issues for which the literature offers competing or contradictory findings due to differences in, for example, research approach, methods, or data sources (“current conflicts”).

3.1 Stabilization Programs Impact on Key Outcomes

The first research question (R1) focused on whether USAID stabilization projects, specifically, and stabilization projects more generally, achieved improvements in the key outcomes of security and conflict events; popular support for the government and for anti-government elements (AGE); community cohesion and resilience; health of the Afghan people; and, economic well-being of the Afghan people. Overall, we found consistent evidence of a small, short-term reduction of violence in some areas and limited evidence for any other effect. We also found some gaps in the existing research regarding the importance of addressing underlying economic conditions when seeking to increase stability in conflict-affected areas and almost no evidence on the types of activities that are needed to ensure the sustainability of key outcomes.

3.1.1 Security

There is substantial evidence in the literature that USAID stabilization programming in Afghanistan had a small, but positive, short-term impact on local security (that is, violence was reduced and/or a higher proportion of the population reported that they felt secure), with considerable regional variation. Evaluations of specific programs provide evidence of micro-level impacts, largely in terms of changes in popular perceptions of their local security environment. The robustness of this evidence is mixed, ranging from the fully representative random probability surveys of the Measuring Impact of Stabilization Initiatives (MISTI) program to qualitative case studies to more anecdotal reviews of existing programs.

However, as is the case in assessing nearly all outcomes of stabilization programming, there are significant gaps in our understanding of how sustainable these impacts are in the medium- to long-term. The academic literature on stabilization aid highlights the ambiguity in the effects of development assistance – which is typically more long-term both in execution and in impact – on conflict and security. Recent empirical analyses suggest development assistance may exacerbate or prolong civil conflict, either by incentivizing insurgent groups to employ greater violence to derail projects that may weaken their position or by increasing all combatants’ uncertainty about the other side’s relative strength (Narang, 2015).

The literature indicates the impact of aid on security is highly dependent on levels of government control and insurgent presence in the districts where projects are implemented: stabilization aid only reduces violence when administered in districts already controlled by pro-government forces (Fishstein and Wilder, 2012; Sexton, 2015). The fact that physical security itself is a key *determinant* to successful program implementation and sustainability makes it difficult to assess the impact of development aid on security as an *outcome*. (Derleth and Alexander, 2010; GIRoA, 2010; MSI, 2013 (2)). This is further complicated by the integration of potential insurgents in local communities. Indeed, broader studies on programs with scope beyond stabilization (e.g., development programs, long-term capacity building initiatives) find that humanitarian assistance

in conflict settings does not have uniform effects and the impact of violence on changes in civilian attitudes depends on whether the perpetrator is viewed as part of their in-group (Lyall, 2016; Lyall, Blair and Imai, 2013).

3.1.2 Support for Government and Anti-Government Elements

We find largely inconsistent evidence on the relationship between stabilization programs and support for either government or anti-government elements. In many cases, the degree to which these programs influenced attitudes was driven by activities outside of the programs' control (see, for example, Altai, 2012 (2) or Altai, 2012 (3)). A key factor in how programs related to attitudinal changes was the degree to which projects were implicated in government corruption. This evidence feeds the broader set of assumptions that the fundamental conflict drivers in Afghanistan are inherently political in nature (e.g., ethnic grievances, inter- and intra-tribal disputes). It is clear that a considerable proportion of Afghan citizens believe the main cause of insecurity to be their own government, which is perceived to be massively corrupt, predatory and unjust. Stabilization programs that rely on using aid to win the population over to such a negatively perceived government face an uphill struggle (Carter, 2013; Fishstein, 2012). There is not sufficient evidence to justify the claim that the Taliban or other AGE were perceived as more effective in addressing the people's highest priority needs of security and access to justice. The evidence is conflicting as to whether individuals viewed support for the Afghan government or the Taliban as zero sum and the extent to which such support can be won through the actions of external actors (Wilton Park Conference, 2010).

3.1.3 Community Cohesion and Resilience

The literature was surprisingly sparse in explicitly defining or assessing social cohesion. Some studies explore this topic by defining impact of conflict on social cohesion through measuring disrupting social engagements and activities (e.g., AIR, 2013; Counterpart, 2005). The degree to which social and community cohesion is important for well-being and stability remains an important gap in the existing research.

Based largely on a handful of project evaluations (e.g., MISTI, ASI), the literature highlights the degree to which there is regional variation in the impact of stabilization programs on community cohesion and resilience. Social capital and local leader satisfaction indices from the final wave of the MISTI evaluation survey indicate perceptions of resilience are strongest in southern districts targeted by the Stability in Key Areas-South (SIKA-S) project, where respondents are most likely to say their community is able to work together to solve problems that come from outside their village. Respondents in SIKA-S districts are also most likely to believe the interests of ordinary people and the interests of women are considered when local leaders make decisions that affect their village/neighborhood. Although Kandahar Food Zone (KFZ) districts are also in the south, those living in KFZ districts perceive the lowest levels of community resilience and

cohesion. Since KFZ districts were selected for inclusion in USAID stabilization programming because of high rates of poppy cultivation, the corrosive effects of the drug trade may explain some of the lack of community resilience and cohesion (MISTI, 2015 (3)). The analyses do not explicitly rule out, however, that differences in reported cohesion could be due to differences in specific aspects of the projects themselves rather than geographic or other sources of variation in the people responding to the surveys.

3.1.4 Health and Economic Well-Being

The very limited evidence on the effect of stabilization programs on health and economic outcomes remains a significant gap in the literature. In part, this is because there were few studies which included economic and/or health indicators explicitly, as shown in Table 2. A handful of other studies, listed below, discuss health outcomes but did not include explicit indicators to evaluate health improvements. A number of studies focused on improvement in health services (AIR, 2013; Beath, Fontini and Enikolopov, 2013 (1a); Child, 2014; MSI, 2013 (2)), but contained limited evidence on whether improvements in access were sustained and on the extent to which such access improved actual health outcomes.

Regarding economic outcomes, there is limited evidence of slight improvements in economic conditions during the implementation of stabilization programs (Altai, 2012 (4), Beath, Fontini and Enikolopov, 2013 (1a); MSI, 2013 (2)). In most cases, this appears to be driven by the direct creation of jobs (Felbab-Brown, 2012; Altai, 2012 (2)), but in some cases job-training programs also had a positive impact (IMU, 2015 (1)). The literature remains conflicted on whether this focus on economic outcomes is desirable. While some research suggests that lack of economic opportunity is a source of some frustration among the general population and ensuring stable economic conditions for Afghans is a critical prerequisite for stabilization (MFA Denmark, 2012; Social Impact, 2016; MSI, 2013 (2)), others argue that focusing on economic conditions distracts attention from the political and social issues that are fundamental to generating grievances and driving conflict (Ellwood, 2013; Gordon, 2011). Absent empirical evidence linking economic conditions to instability and conflict in these settings, it remains ambiguous whether addressing economic conditions should be a necessary element of stabilization programs.

3.2 Time Horizon of Effects

The second research question (R2) was focused on the time horizon over which any effects were apparent. In many cases, programs focused on generating rapid effects within a very short window (3-6 months). Other programs were focused on the medium-term (6-18 months) or long-term (18+ months).⁹ Some stabilization programs also considered the impact on key indicators

⁹ We note that these time horizons are only medium and long term in the context of “stabilization” programs.

over an even longer timeframe--3-5 years from implementation. Very few impacts were seen over this longer time horizon for two reasons. First, and unavoidably in conflict-affected settings, the dynamic nature of the environment makes it hard to find any effects from programming after more than a year. Too many other inhibiting and amplifying factors are changing along with independent but confounding drivers of instability. In such conditions, the signal to noise ratio is typically quite low after more than six months.¹⁰ Second, program outcomes are not typically designed to be measured years after implementation is complete; programs do not often include this in their budgets. We also did not find retrospective studies on the effects of major programs after program implementation had concluded.

3.2.1 Limited Evidence of Short-Term Impact

Nearly all well-designed studies, including experimental and quasi-experimental quantitative approaches (e.g., Beath, Fontini and Enikolopov, 2013 (1b); MSI, 2014 (7)), government-initiated evaluations (Bohnke, Koehler and Zurcher, 2014; Bohnke and Zurcher, 2013; Norad, 2012), qualitative reviews (Fishstein, 2012), and historical accounts (Goodhand, 2002) consistently indicated that any stabilization program effects (whether positive or negative) were short-term at best. This appears to be true of both US and international civilian-led programs, such as the ASI programs (Altai, 2012 (1)) or NSP (Beath, Fontini and Enikolopov, 2013 (1b)), and military-led programs, such as CERP (Chou, 2012). It also applies to the stabilization efforts of non-USG foreign donors, such as Norway (Norad, 2012), Germany (BMZ, 2010), and the UK (DFID, 2009). While several of these programs generated shifts in the security environment, government capacity or reported attitudes, there is no evidence that these shifts were sustained after the programs concluded.

The literature is divided on the desirability of short-term or long-term stabilization aid. On the one hand, some reports emphasize that a focus on short-term objectives is essential to help the host nation get off life support and on a sustainable path to recovery (Cole and Hsu, 2009; Narang, 2015; SFRC, 2011). Proponents of a short-term focus argue that quick-impact gains address specific, pressing needs and build the foundation for other, longer-term activities (by either the Afghan government or other international actors). The critique of a short-term focus centers on the assertion that rapid gains or “quick wins” result in outcomes that are unsustainable without continued foreign support, ultimately breeding dependency and resentment among the Afghan population (Brown, 2014; Felbab-Brown, 2012; ICG, 2011; Miakhel, 2010). Specifically, Felbab-Brown (2012) notes that the focus on short-term gains does not address structural drivers of instability (especially in rural Afghanistan) and, as a result, quick-impact

Development programs, for example, would operate and expect impact over much longer time horizons.

¹⁰ In impact evaluations in conflict zones, this is a general problem for most any outcome other than basic demographics among geographically stable populations.

oriented programs, such as CERP-funded projects, have tended to replace government capacity rather than grow it. Taylor (2010), in particular, argues that the donor community should shift its focus from “quick wins” to sustainability. While logically appealing, there is limited evidence that empirically validates this assertion.

3.2.2 Gaps in Understanding Long-Term Impact

While the short-term and transitory nature of stabilization program impacts appears to be well-documented, at least for some key indicators, their longer-term impacts are less understood. The research on lack of long-term outcomes largely focuses on the reasons for lack of sustainability. These reasons include: projects were explicitly designed to achieve short-term stability, rather than long-term sustainability (Carter, 2013; Cole and Hsu, 2009); many projects had unrealistic goals and mechanisms to facilitate sustainable outcomes (DFID, 2009; Felbab-Brown, 2012); some projects were not focused on key drivers or issues relevant for change (Brown, 2014); and specific factors that had pervasive effects on program implementation (e.g., corruption, personnel turnover) created barriers for sustainable outcomes (DFID, 2009; Gordon, 2014; ICG, 2011; Miakhel, 2010; MSI, 2015 (1)).

In addition to the programmatic reasons for lack of longer-term and sustainable impact, most efforts for monitoring and evaluation were focused on oversight and implementation and did not continue to assess outcome changes after program completion. This was especially true for the experimental and quasi-experimental evaluations of stabilization programming in our review. In large part, this is because assessing outcome changes after program completion is costly and it risks the inclusion of other confounding factors, which could make any results difficult to interpret. The environment in Afghanistan specifically, and many conflict-affected areas generally, is dynamic given the social turmoil and many government and international activities that may affect measurable indicators. When evaluating an individual program many years after its completion, other programs and sources of turmoil can confound and bias estimates of the impact of the specific program in question. This lack of measurement also impacted the design and evaluation of longer-term programs conducted with more traditional development objectives (UNDP, 2014).

3.3 Impact of Military Presence

Given the prominence of the military and its expanded role in stabilization program operation and execution, it is critical to consider the interactive effects between the military and civilian activities. In particular, we focused on how the presence of the military impacted the outcomes of stabilization projects (R3). On the whole, the literature suggests that the military played an important role in providing baseline levels of security to facilitate program operation and that some of the military’s programs—namely CERP—were important stabilization programs in their

own right. However, the differences in objectives, timelines, and cultures resulted in inefficiencies and sometimes limited the effectiveness of operations. At a more strategic level, there remains an active, ongoing debate on the degree to which assistance programs should be supported or executed by the military.

3.3.1 International Military Forces have a Role in the Basic Execution of Stabilization Programs

At a purely tactical level, much of the literature recognizes the importance of military presence during program execution to assist in providing the basic level of security needed for program execution (DoD JCOA, 2006; Felbab-Brown, 2012; ICG, 2011; Kapstein and Kathuria, 2012; Sexton, 2015; Taylor, 2010). Absent this support (and sometimes even with it), the security situation inhibited even basic tasks needed for program operation (see, for example, Altai, 2012 (5)). This vital function was acknowledged even among those critical of the military's role in the stabilization context. As noted by Taylor (2010): "security is still the major issue inhibiting project implementation in stabilization contexts. Donors need to find more innovative, effective and varied ways to deal with security issues in aid delivery." Thus while the military may have been instrumentally useful in allowing basic operations, that support did not negate the potential inhibiting effect of poor security.

3.3.2 Civilian and Military Implementers Have Conflicting Objectives

Although the military played an important and well-recognized role in supporting stabilization efforts, the differences between civilian and military objectives (Gordon, 2014), timeframes (Davids, Rietjens and Soeters, 2010), and culture (Altai, 2012 (4)) resulted in a range of inefficiencies and conflicting activities. While the empirical evidence of problems resulting from a lack of civil-military cooperation is largely qualitative, it is robust, widespread, and consistent.

A range of studies consistently identified these issues and noted the military's role in affecting successful outcomes. In particular, ICG (2011) noted that "in their haste to demonstrate progress, donors have pegged much aid to short-term military objectives and timeframes. As the drawdown begins, donor funding and civilian personnel presence, mirroring the military's withdrawal schedule, may rapidly decline, undermining oversight and the sustainability of whatever reconstruction and development achievements there have been." Moreover, many strategic plans and policy documents did not fully recognize different objectives pursued by civilian representatives and the military leadership, resulting in uncoordinated and sometimes conflicting efforts (Viehe, Afshar and Heela, 2015).

3.3.3 The Securitization of Development Aid is a Controversial Concept

An important and ongoing conflict in the literature is the degree to which the impact of the military on stabilization outcomes should be analyzed in the context of a wider debate over the

securitization of development aid. While much of the evidence related to this controversy stems from papers that lack rigorous research designs, their anecdotal and experience-informed insights do center on a common theme. According to one of the first academic studies of the role of international aid in Afghanistan, “maximalists” argue aid should be consciously used as an instrument of peace building, while “humanitarian minimalists” contend the implementation of aid to support military objectives leads to a distortion of traditional mandates, especially neutrality and impartiality (Goodhand, 2002). The implementation of the Provincial Reconstruction Team (PRT) framework highlighted both the potential benefits and implicit tensions in civilian-military development collaboration (McNerney, 2006; Waldman, 2008).

More broadly, Howell and Lind (2009) note that “the convergence of military and development objectives and the subordination of the latter to the former has co-opted civil society into stabilization and state-building strategies in Afghanistan as a way of strengthening the state and has undermined the legitimacy of civil society and contributed to negative popular attitudes toward NGOs.” While these concerns directly contradict the stated need for the military to serve a security function for program execution, there is limited evidence to either support or reject the notion that securitizing development assistance is problematic.

3.4 Synergies and Confounding Factors with Other Donors

Motivated in part by the degree to which the military presence affected stabilization programs, we next reviewed the evidence on the synergies and confounding factors that may exist between stabilization programs by different actors (other parts of the USG, other countries, Afghan government, international organizations like the World Bank, etc.) (R4). There is very limited evidence in the literature on what types of synergies beyond the civil-military partnership were effective. However, several articles highlighted the importance of coordination between different countries and the host nation (e.g., Afghanistan) in ensuring success (Ellwood, 2013; MSI, 2014 (2); Altai, 2012 (7); Goodhand, 2002; Gordon, 2012; Taylor, 2010; Viehe, Afshar and Heela, 2015; Waldman, 2008). Ellwood (2013) summarized this by noting that “despite the scale of international aid that has been poured into the country (estimated to be some \$430 billion), conflicting agendas, poor coordination, lack of overall ownership, an absence of regional economic strategies, and an ignorance of local requirements have led to time, effort, and finances wasted on an industrial scale.” Despite the repeated observations on these issues, there is limited evidence on the specific ways to design or ensure international donor and/or implementer coordination that would enable or inhibit success. As noted in greater detail in the recommendation section, this issue could be addressed by future monitoring and evaluation efforts, which would focus on identifying programmatic features as well as overall impact.

3.5 Impact of Specific Project Aspects

We next turn to the programmatic features that could amplify or inhibit the impact of stabilization programs, including the size and type of contract used to facilitate implementation, as well as the sectors (agriculture, infrastructure, skills, etc.) in which projects operate (R5).

3.5.1 Small-Scale Programs Produced Quick Impact Results, but Did Not Address Drivers of Instability

A number of studies (Altai, 2012 (7); Chou, 2012; Child, 2014; Goodhand, 2002; Gordon, 2011; Kapstein and Kathuria, 2012; Nagl, Exum and Humayun, 2009) highlight the intuition that smaller projects can be targeted at important, specific gaps and seem less likely to fuel instability. We note that these small projects are small in the scale of any individual project; the programs which fund such projects may be quite large including in many instances nationwide in scope. Small projects have a variety of beneficial features: they are easier to manage; they are less likely than large infrastructure projects to attract attention from corrupt officials or to become targets for enemy sabotage; and outputs are small and less likely to become a source of conflict. For each of three reconstruction programs included in our analysis (NSP, LGCD, CERP), project spending was not associated with statistically significant reductions in violence. The one exception was small-scale development aid that was conditional on information sharing by the community; this incentivized approach did appear to be somewhat effective in reducing violence (Chou, 2012). The literature does not provide evidence of increasing returns from a cost-effectiveness perspective; small projects do not have a different impact on outcomes, such as violence or support for the government, relative to larger-scale projects (see Child, 2014). Additionally, Afghans reacted positively to large-scale programs that are populated by small, community driven projects such as NSP because the funding dispersed was too small to be siphoned off by powerful interests while still providing meaningful benefits to communities (Gordon, 2011). This is important for minimizing the risk of unintended negative outcomes as well. For instance, any corruption that affects community-driven, local projects is by its nature smaller in scale and thus less likely to delegitimize the national government (Kapstein and Kathuria, 2012).

However, Felbab-Brown (2012) noted an important caveat: small-scale stabilization programs do not address the structural deficiencies of the rural economy in Afghanistan, such as inadequate infrastructure, lack of economic opportunities, and motivations for many lower-level grievances. Moreover, small-scale programs tended mostly to replace government capacity rather than to grow it, further exacerbating the periphery-center divide. As such, while small-scale projects may be effective at impacting short-term outcomes, they are unlikely to have more permanent, sustainable impacts.

3.5.2 Large-Scale Programs Created Unrealistic Expectations and Experienced Higher Levels of Corruption and Targeting

Larger programs – often related to construction – were problematic for several reasons and were less likely to have significant impacts. Based on the Afghanistan experience, stabilization programming is unlikely to generate large or persistent effects. From well-designed experimental and quasi-experimental approaches (e.g. Beath, Fontini and Enikolopov, 2013 (1b); MSI, 2014 (7)) to government-initiated qualitative reviews (Bohnke, Koehler, Zurcher, 2014; Bohnke, Zurcher, 2013; Norad, 2012) to historical accounts (Goodhand, 2002), the evidence consistently indicates stabilization programming has small, generally transitory, impacts (either positive or negative). Programs such as CERP (Chou, 2012; Child, 2012; Sexton 2015) or some of the ASI programs (Altai, 2012 (1)) that have been “successful” do not generate large shifts in security, attitudes, or capacity, though they may have short-term, positive impacts. This is relevant for both managing the expectations of what stabilization programs may accomplish and for considering how to design measurement and evaluation efforts to detect relatively small effects. For example, expectations of positive program outcomes among the Afghan population, when unmet, resulted in disappointment and disillusionment when these programs, in their view, did not deliver on their promises (Gordon, 2012). That disappointment could underlie the continued dissatisfaction among Afghan people with international forces and/or the Afghan Government, making it increasingly difficult for future programs to gain local support.

Second, large programs appeared to be much more susceptible than their smaller counterparts to negative forces, such as corruption and violence. This was especially true in the construction sector, which Afghans tended to view as low quality, corrupt and highly criminalized (Gordon, 2012; Sud, 2013). Moreover, these large projects were subject to criminal and insurgent targeting for violence and other attacks (SIGAR, 2011 (1)). Many large contracts also involved subcontracting that most Afghan respondents regarded simply as a legalized form of corruption (Gordon, 2012).

3.6 Summary of Commonalities across Successful or Unsuccessful Projects

Based on the existing evidence, there were a number of commonalities that we identified when looking across a variety of successful or unsuccessful stabilization projects at different times, with different implementers, in a range of different sectors (R6). Many of these commonalities relate to external conditions or factors which influence success. Thus, rather than focusing on which underlying drivers of instability (e.g., ethnic tension, poverty) may have been most relevant for the specific setting, we find the literature is most instructive in identifying relevant enabling and inhibiting factors (e.g. political will), regardless of the specific driver or source of instability. These factors include: host nation coordination and commitment, limiting the extent of corrosive corruption, ensuring baseline levels of security to facilitate basic implementation

and oversight, and ensuring appropriate staffing, both in terms of skills and in terms of longevity of deployment.

We note an important caveat in this analysis: in many cases the absence of such key factors is itself what drives instability; but, of course, not in all cases. As noted by Mikulaschek and Shapiro (2016): "...there is no reason to expect the same correlation between a given cause (e.g., poverty) and both the onset of conflict and sub-national variation in its intensity." Thus the presence or absence of these factors can occur in areas with varying degrees of stabilization and may directly affect the programmatic success as well as the overall level of stabilization in a given area. This is relevant in the Afghan context – as well as when applying these findings to other conflict settings – in managing expectations on what may and may not work.

3.6.1 The Afghan National Government's Commitment to Reform is a Fundamental Prerequisite for Success

As noted above, coordination with the host-nation is critical for effective implementation. However, beyond simple coordination, a real commitment to building capacity and reforming ineffective processes is critical to building a responsive, legitimate government (Cole and Hsu, 2009). A regularly noted limiting factor was the lack of capacity and willingness to reform by the Afghan government at both local and national levels.

3.6.2 Impact of Corruption is Pervasive and Corrosive

Not surprisingly, a wide variety of studies noted that corruption was a key--if not the single most important--issue affecting support for the Afghan government, support for insurgents, and attitudes towards foreign forces (AIR, 2013; Altai, 2012 (1); Altai, 2012 (3); Ellwood, 2013; Felbab-Brown, 2012; Fishstein, 2012; Gordon, 2012; Miakhel, 2010; MSI, 2013 (2); Viehe, Afshar and Heela, 2015). As Fishstein (2012) summarized: "while respondents...did report some short-term benefits of aid projects, it appears that corruption, tribal politics, and the heavy-handed behavior of international forces neutralized whatever positive effects aid projects might have delivered." Gordon (2012) similarly noted that Afghans consistently described development projects negatively; not only were projects failing to build support for government among civilians but they were increasing perceptions of corruption and distrust in government.

3.6.3 Security is a Key Determinant of Program Success and Sustainability

As noted in Section 3.3, the military plays a key role in providing basic security for program implementation and execution. More broadly, security is an important factor in creating the preconditions for program success (AIR, 2013; Davids, Rietjens and Soeters, 2010; IMU, 2015 (2)). This is distinct from suggesting security itself is sufficient to produce programmatic success, as noted in a number of critiques on the securitization of development assistance.

Nevertheless, most Afghans, when surveyed, noted security as an important factor when considering the effectiveness of development projects in their communities (BMZ, 2010). However, in many cases, even though those surveyed had not themselves experienced violence, the perception of violence was an important determinant of attitudes towards the Afghan government (IMU, 2015 (4)). Thus, when planning for programs, attention must be paid to not only how to establish security in an area but also what factors affect public perceptions of security among the local population.

3.6.4 Employing Competent, Long-term, and, Ideally, Local Staff are Essential Elements of Successful Projects

Even the best designed programs cannot be effectively implemented and achieve the desired impact without appropriate staffing. A range of audits and evaluations found consistent, substantial negative effects on program effectiveness due to lack of adequate staff. This included understaffing for key programs (DFID, 2009), lack of appropriately trained staff (DoD JCOA, 2006), rapid turnover of staff (MSI, 2014 (2)), restrictions on the mobility of staff (OIG, 2015), and the underuse of capable local staff (Miakhel, 2010). As a result, programs were often not implemented as intended and could not be modified and adapted to meet requirements. Moreover, absent sufficient staff in the field, stabilization programs lacked the appropriate oversight and feedback to even identify when such modifications might be necessary. Programs in which staff were in the field and accessible, especially smaller programs such as OTI's CCI programs, were better able to achieve modest program goals (IMU, 2015 (3)).

4. CONCLUSIONS AND RECOMMENDATIONS

The overall findings of this report can be summarized as follows: manage expectations because big changes in any outcomes are unlikely, smaller programs are more likely to achieve modest outcomes, security and corruption must be addressed to enable success, and appropriate staffing is critical for implementation.

These findings are supported by a systematic compilation of evidence from the available literature through late-2016. We must highlight the fact that almost none of these key points are new. Although many of the studies included in this analysis noted key points from evidentiary base at the time that are reflected in our report as well—few of those points were heeded after publication. Many of our findings about the impact of development programming based on studies through 2016 were made in ICG (2011), which argued that: “The impact of international assistance will remain limited unless donors, particularly the largest, the U.S., stop subordinating programming to counter-insurgency objectives, devise better mechanisms to monitor implementation, adequately address corruption and wastage of aid funds, and ensure that recipient communities identify needs and shape assistance policies.”

Of particular note is the repeated, widespread recommendation for improved monitoring and evaluation in order to improve future program performance. This recommendation is succinctly summarized by Bohnke, Koehler and Zurcher (2014) as follows: “...if the international community is serious about rigorous impact evaluations, it must pressure donors and implementing actors for much higher standards for recording and sharing data!” Improved monitoring, evaluation, and learning could be achieved through adopting the improvements and changes suggested in Department of State 2011, including addressing key gaps such as: “USAID does not follow a uniform approach to the conduct of evaluations;” “the number of evaluations conducted by USAID is very small;” “most interventions are not evaluated;” and “although USAID mandates that each major intervention should be evaluated at least once, the mandate appears not to have been followed.” Key improvements that were recommended at that time include the need for more detailed statements of work to ensure appropriate, feasible data collection and evaluation plans; the importance of methodologically sound and clearly documented evaluation designs; and clear, explicit and publicly available presentation of data, findings, and recommendations from these evaluations. Should a future contingency entail a large, international, multi-year stabilization effort, programs and evaluations should be designed, implemented and modified to take into account these recommendations to improve the chances for success.

REFERENCES

Program Evaluations

- AIR Consulting for Creative Associates International. *Assessment Methodology and Pilot Baseline Studies for Four CCI Districts: KC 6, 7, 8; KC 9; Zhari, and Muqur*, February, 2013.
- Altai Consulting. *Afghanistan Stabilization Initiative (ASI) Third Party M&E and Strategic Support: Barmal-Shkin District Report*, February, 2012.
- *Afghanistan Stabilization Initiative (ASI) Third Party M&E and Strategic Support: Khas Uruzgan District Report*, February, 2012.
- *Afghanistan Stabilization Initiative (ASI) Third Party M&E and Strategic Support: Khogyani District Report*, February, 2012.
- *Afghanistan Stabilization Initiative (ASI) Third Party M&E and Strategic Support: Sar Hawza District Report*, February, 2012.
- *Afghanistan Stabilization Initiative (ASI) Third Party M&E and Strategic Support: Sarkani District Report*, February, 2012.
- *Afghanistan Stabilization Initiative (ASI) Third Party M&E and Strategic Support: Sayredabad District Report*, February, 2012.
- *Afghanistan Stabilization Initiative (ASI) Third Party M&E and Strategic Support: Urgun District Report*, February, 2012.
- Caerus Analytics. *CCI Midterm Performance Evaluation Report*. May 2014.
- Independent Monitoring Unit, Rahman Safi Consulting. *IMU Case Study: GIROA Perspectives on OTI/CCI Programming*, November, 2015.
- *IMU Case Study: Winter Preparedness Packages*, December, 2015.
- *IMU Case Study: The Effectiveness of Infrastructure Projects in Increasing Linkages Between Local Populations and Government*. (Pending receipt of copy)
- *IMU Case Study: Youth Programming in South, East, North and West Afghanistan*. (Pending receipt of copy)
- Management Systems International. *ACAP II End-Line Report: Examining ACAP II's PMP Indicators*, July, 2014.
- *ACAP II Final Performance Evaluation*, February, 2015.
- *Kandahar Food Zone Mid term Performance Evaluation*, March, 2015.
- *MISTI Desk Review of Stabilization Resources and References*, July, 2012.
- *MISTI Stability Trends and Impact Evaluation Survey: Design, Methods, Pilot and Indicators*, November, 2012.
- *Measuring Impacts of Stabilization Initiatives (MISTI): Baseline Stabilization Trends and Impact Evaluation Survey (Presentation)*, February, 2013.
- *MISTI Stabilization Trends and Impact Evaluation Survey Analytical Report, Wave 1 (Baseline): Sep - Dec 2012*, July, 2013.

----- *MISTI Stabilization Trends and Impact Evaluation Survey Analytical Report, Wave 2: May 18 - Aug 7, 2013*, April, 2014.

----- *MISTI Stabilization Trends and Impact Evaluation Survey Analytical Report, Wave 3: Nov 16, 2013 - Jan 30, 2014*, July, 2014.

----- *MISTI Stabilization Trends and Impact Evaluation Survey Analytical Report, Wave 4: Apr 28 - Jun 30, 2014*, December, 2014.

----- *MISTI Stabilization Trends and Impact Evaluation Survey Analytical Report, Wave 5: Sep 28 - Nov 3, 2014*, November, 2015.

----- *SIKA-North Mid-Term Performance Evaluation*, August, 2014.

----- *SIKA-South Mid-Term Performance Evaluation*, October, 2014.

----- *SIKA-West Mid-Term Performance Evaluation*, March, 2014.

RAND. *Peer Review of the MISTI Survey and Evaluation Methodology*, September, 2014.

Social Impact. *Final Performance Evaluation of the USAID/OTI Community Cohesion Initiative in Afghanistan*, February, 2016.

Government Documents

Bennett, J., Alexander, J., Saltmarshe, D., Phillipson, R., and Marsden, P., *Evaluation of DFID's Country Programmes: Afghanistan 2002-2007*, UK Department for International Development, May 2009.

Evaluation Department, Norwegian Agency for Development Cooperation. *Evaluation of Norwegian Development Cooperation with Afghanistan 2001-2011*. 2012.

Federal Ministry for Economic Cooperation and Development, Government of Germany. *Assessing the Impact of Development Cooperation in North East Afghanistan 2005 – 2009*. Undated.

Government of the Islamic Republic of Afghanistan, Baawar Consulting Group. *Joint Evaluation of the Paris Declaration Phase 2: Islamic Republic of Afghanistan*. 2010.

Independent Commission for Aid Impact. *DFID: Programme Controls and Assurance in Afghanistan*, March, 2012.

Ministry of Foreign Affairs, Government of Denmark. *Evaluation of Danish Development Support to Afghanistan*, August, 2012.

Office of the Director of Foreign Assistance, Department of State. *A Meta-Evaluation of Foreign Assistance Evaluations*, June, 2011.

Office of the Inspector General. *Audit of USAID/Afghanistan's Strategy for Monitoring and Evaluating Programs Throughout Afghanistan*, December, 2015.

Senate Foreign Relations Committee Majority Staff. *Evaluating U.S. Foreign Assistance to Afghanistan: A Majority Staff Report Prepared for the Use of the Committee on Foreign Relations United States Senate*, June, 2011.

Special Inspector General for Afghanistan Reconstruction. *Afghanistan's National Solidarity Program Has Reached Thousands of Afghan Communities, but Faces Challenges that*

- *Could Limit Outcomes*, March, 2011.
- *Commander's Emergency Response Program in Laghman Province Provided Some Benefits, but Oversight Weakness and Sustainment Concerns Led to Questionable Outcomes and Potential Waste*, January, 2011.
- *Progress Made Toward Increased Stability under USAID's Afghanistan Stabilization Initiative-East Program but Transition to Long Term Development Efforts Not Yet Achieved*, June, 2012.
- *USAID Spent Almost \$400 Million on an Afghan Stabilization Project despite Uncertain Results, but Has Taken Steps to Better Assess Similar Efforts*, April, 2012.
- Sud, I., *Afghanistan: A Synthesis Paper of Lessons from Ten Years of Aid*. Independent Evaluation Group of the World Bank, February, 2013.
- UKAID Stabilisation Unit. *Stabilisation Case Study: Infrastructure in Helmand, Afghanistan*, November, 2010.
- UNDP Independent Evaluation Office. *Assessment of Development Results: Afghanistan*, July, 2014.

Academic Literature

- Beath, A., Fontini, C., and Enikolopov, R. *Randomized Impact Evaluation of Afghanistan's National Solidarity Programme*, July 2013.
- , *Winning Hearts and Minds through Development: Evidence from a Field Experiment in Afghanistan*, MIT Political Science Department Research Paper No. 2011-14, April, 2016.
- Böhnke, J. R., and Zürcher, C. "Aid, minds and hearts: The impact of aid in conflict zones," *Conflict Management and Peace Science*, 30(5), August, 2013.
- Böhnke, J. R., Koehler, J., and Zürcher, C. "Assessing development cooperation in northeast Afghanistan with repeated mixed-method surveys," in Andersen, O. W., Bull, B., and Kennedy-Chouane, M., eds., *Evaluation Methodologies for Aid in Conflict*, New York: Routledge, 2014.
- Carbonnier, G. "Humanitarian and Development Aid in the Context of Stabilization: Blurring the Lines and Broadening the Gap," in Muggah, R., ed. *Stabilization Operations, Security and Development: States of Fragility*, New York: Routledge, 2014.
- Carter, W.R. "War, Peace and Stabilisation: Critically Reconceptualising Stability in Southern Afghanistan," *Stability: International Journal of Security & Development*, 2 (1): 15, June, 2013.
- Child, T. "Hearts and minds cannot be bought: Ineffective reconstruction in Afghanistan," *The Economics of Peace and Security Journal*, 9:2, May 2014.
- Chou, T. "Does development assistance reduce violence? Evidence from Afghanistan," *The Economics of Peace and Security Journal*, 7:2, August, 2012.
- Crost, B., and Johnson, P. *Aid Under Fire: Development Projects and Civil Conflict*, Harvard

- Kennedy School, Belfer Center for Science and International Affairs Discussion Paper 2010-18, November 2010.
- Davids, C., Rietjens, S., and Soeters, J. "Measuring Progress in Reconstructing Afghanistan," *Baltic Security and Defence Review*, 12:1, January, 2010.
- Dennys, C. "For Stabilization," *Stability: International Journal of Security & Development*, 2 (1): 1, February, 2013.
- Goodhand, J., and Sedra, M. "Who owns the peace? Aid, reconstruction, and peacebuilding in Afghanistan," *Disasters*, 34 (S1), January, 2010.
- Goodhand, J., and Sedra, M., eds. *The Afghan Conundrum: intervention, statebuilding and resistance*, New York: Routledge, 2015.
- Gordon, S. "Afghanistan's Stabilization Program: Hope in a Dystopian Sea?" in Muggah, R., ed. *Stabilization Operations, Security and Development: States of Fragility*, New York: Routledge, 2014.
- Howell, J. and Lind, J. "Manufacturing Civil Society and the Limits of Legitimacy: Aid, Security and Civil Society after 9/11 in Afghanistan," *European Journal of Development Research*, 21, 2009.
- Lyall, J. "Civilian Casualties and the Conditional Effects of Humanitarian Aid in Wartime," May, 2016.
- Lyall, J., Blair, G., and Imai, K. "Explaining Civilian Attitudes in Wartime: A Survey Experiment in Afghanistan," *American Political Science Review*, 107:4, November, 2013.
- Mac Ginty, R. "Against Stabilization," *Stability: International Journal of Security & Development*, 1 (1): 20-30, November, 2012.
- McNerney, M. "Stabilization and Reconstruction in Afghanistan: Are PRTs a Model or a Muddle?" *Parameters*, Winter 2005-06.
- Mikulaschek, C., and Shapiro, J. "Lessons on political violence from America's post-9/11 Wars," March, 2014.
- Sexton, R. "Aid as a Tool Against Insurgency: Evidence from Contested and Controlled Territory in Afghanistan," *American Political Science Review*, Nov 2015.

Policy/Think-Tank Studies

- Agoglia, J., Dziedzic, M., and Sotirin, B., eds., *Measuring Progress in Conflict Environments (MPICE): A Metrics Framework for Assessing Conflict Transformation and Stabilization*. Washington, DC: USIP Press, June, 2010.
- Brown, F. *Rethinking Afghan Local Governance Aid After Transition*, USIP Special Report 349. August, 2014.
- Cole, B., and Hsu, E. *Guiding Principles for Stabilization and Reconstruction*, Washington, DC: USIP Press and US Army Peacekeeping and Stability Operations Institute, 2009.

- Derleth, J., and Alexander, J. "Stability Operations: From Policy to Practice," *Prism*, 2:3, 2010.
- van Stolk, C., Ling, T., Reding, A., and Bassford, M. *Monitoring and Evaluation in Stabilisation Interventions*, Cambridge: RAND Europe, 2011.
- Ellwood, T. *Stabilizing Afghanistan: Proposals for Improving Security, Governance and Aid/Economic Development*, Atlantic Council, April, 2013.
- Felbab-Brown, V. "Slip-Sliding on a Yellow Brick Road: Stabilization Efforts in Afghanistan," *Stability: International Journal of Security & Development*, 1 (1): 4-19, November, 2012.
- Fishstein, P., and Wilder, A. *Winning Hearts and Minds? Examining the Relationship between Aid and Security in Afghanistan*, Tufts University Feinstein International Center, January, 2012.
- Gordon, S. *Winning Hearts and Minds? Examining the Relationship between Aid and Security in Afghanistan's Helmand Province*, Tufts University Feinstein International Center, April, 2011.
- International Crisis Group, *Aid and Conflict in Afghanistan*, Asia Report No. 210, August, 2011.
- Kapstein, E., and Kathuria, K. *Economic Assistance in Conflict Zones: Lessons from Afghanistan*, Center for Global Development, October, 2012.
- Miakhel, S. *A Plan to Stabilize Afghanistan*, Centre for International Governance Innovation, May, 2010.
- Nagl, J., Exum, A., and Humayun, A. *A Pathway to Success in Afghanistan: The National Solidarity Program*, Center for a New American Security Policy Brief, March, 2009.
- Open Society Foundations, *The Trust Deficit: The Impact of Local Perceptions on Policy in Afghanistan*, OSF Regional Policy Initiative on Afghanistan and Pakistan Policy Brief No. 2, October, 2010.
- Taylor, M. "Civilian-Military Cooperation in Achieving Aid Effectiveness: Lessons from Recent Stabilization Contexts," *The 2010 Brookings Blum Roundtable Policy Briefs*, April, 2010.
- Thompson, E. *Winning 'Hearts and Minds' in Afghanistan: Assessing the Effectiveness of Development Aid in COIN Operations*, Report on the Wilton Park Conference 1022, March, 2010.
- Viehe, A., Afshar, J., and Heela, T. *Rethinking the Civilian Surge: Lessons from the Provincial Reconstruction Teams in Afghanistan*, Center for American Progress, December, 2015.
- Waldman, M. *Falling Short: Aid Effectiveness in Afghanistan*. ACBAR/OXFAM, March, 2008.