

FROM INPUT TO IMPACT
**EVALUATING TERRORISM
PREVENTION PROGRAMS**

PETER ROMANIUK AND NAUREEN CHOWDHURY FINK



Center on Global
Counterterrorism
Cooperation

FROM INPUT TO IMPACT: EVALUATING TERRORISM PREVENTION PROGRAMS

By Peter Romaniuk and Naureen Chowdhury Fink

A project of the Center on Global Counterterrorism Cooperation

With the generous support of the Governments of
Canada, the United Kingdom, and the Netherlands

© Center on Global Counterterrorism Cooperation, 2012.

www.globalct.org



Ministry of Foreign Affairs of the
Netherlands



Foreign &
Commonwealth
Office



Public Safety
Canada

Sécurité publique
Canada

Table of Contents

Introduction	2
Terrorism Prevention in Context	4
Evaluating Terrorism Prevention Programs: Challenges and Opportunities	8
Evaluation Design	10
Implementing Evaluations	14
Analysis	16
Further Research and Follow-up	18
Suggested Readings	20

Introduction

FOR MUCH OF MODERN HISTORY, THE TASK OF countering terrorism has largely been the domain of security agencies utilizing “hard power” tools, including intelligence and law enforcement, alongside the selective use of Special Forces and the military. Although expertise in the disruption and suppression of terrorism accumulated over time, few states developed and implemented policies aimed specifically at terrorism prevention. Where they had, these measures were conceived relatively narrowly, covering public diplomacy and strategies for managing the media. For many states, the terrorist attacks of 11 September 2001 and subsequent attacks around the world prompted a significant reevaluation of counterterrorism approaches. As a result, the strategies and tactics utilized have broadened. Traditional tools maintain a certain prominence but have been complemented by new and innovative responses to a complex and evolving range of threats posed by terrorists and violent extremists. In the years following 9/11, threat assessments necessarily focused on the dangers posed by al-Qaida, its affiliates, and those inspired by its ideology. There are signs, however, that the threat today is more complex and diffuse, comprising extremists from all parts of the ideological spectrum who may act in small “self-starter” groups or, in some cases, as “lone wolves.” The paths to extremism are more varied than ever before, and as our understanding of contemporary patterns of radicalization has advanced, terrorism prevention initiatives have become more prevalent in the counterterrorism repertoire at the national and multilateral levels.

As many states have elaborated terrorism prevention strategies in recent years, they have begun to confront similar challenges. Among these is the challenge of program evaluation. Is the turn toward prevention an effective response to the diverse extremist threats that states face today? How can the effectiveness of prevention policies be measured? What approaches have states advanced in evaluating the impact of terrorism prevention initiatives? In responding to this challenge, can lessons be gleaned from efforts to evaluate programs in related policy domains?

This policy report provides an initial discussion of these questions. It draws on the discussions during a meeting entitled “Colloquium on Measuring Effectiveness in Counterterrorism Programming” and held in Ottawa on 9–10 February 2012, as well as discussions with experts, government officials, and an initial desktop literature review. The colloquium was developed and organized by the Center on Global Counterterrorism Cooperation with the support of the governments of Canada, the Netherlands, and the United Kingdom. Participants included counterterrorism and terrorism prevention experts from national governments and international organizations, as well as academics and practitioners, in these and related fields. Participants agreed that the emphasis on program evaluation and terrorism prevention is timely and appropriate. Given the responsibility of demonstrating to their citizens the effectiveness of their terrorism prevention measures, governments are especially concerned to know which ones have the most impact.

At the same time, governments are facing budgetary pressures and are keen to optimize limited resources. Consequently, there is a heightened demand for evaluation by many governments seeking to identify effective prevention practices.

In the next section of this report, we put the current discussion about program evaluation and terrorism prevention into context, defining the scope of prevention policies and the range of activities that are pursued toward their implementation. We describe the challenges that states face in evaluating prevention efforts. In some ways, these challenges are similar to those faced in other areas of counterterrorism policy. In other ways, terrorism prevention presents specific dilemmas because the outcomes of preventive measures may be intangible; after all, successful prevention would result in a “nonevent.” This report then surveys state experiences to examine how these challenges have been addressed and the related evaluation practices developed.

We see a preliminary trend toward three broad approaches among states. Some have elaborated a *multidimensional* approach, specifying metrics and gathering evaluation data at multiple levels to assess programmatic and broader impacts. Other states have preferred a *vertical* approach, focusing their initial evaluation activities on their specific prevention interventions. Still other states have opted for a *horizontal* approach, examining efforts across a range of government agencies and programs in support of the objective of terrorism prevention. This typology of approaches may inform future developments in this field. We also look to related policy domains, especially those where the key objective of an intervention is to prevent a particular outcome, to glean lessons learned because some of the challenges facing counterterrorism practitioners as they attempt to evaluate prevention initiatives have been confronted by public, private, and nongovernmental organization—sector actors in the past. The colloquium provided an initial opportunity to “learn by analogy” and draw on those experiences, experts, and expertise.

The report concludes by outlining a series of opportunities to advance the initial findings of the colloquium. In sum, as discussions of terrorism prevention continue at the national level and in multilateral fora, we foresee opportunities to accumulate knowledge and facilitate collaboration toward the goal of enhancing the effectiveness of this innovative tool of counterterrorism policy.

Terrorism Prevention in Context

FOR SEVERAL YEARS PRIOR TO 9/11, LEADING terrorism analysts noted an imbalance in the way many states formulate and implement counterterrorism policy. For example, counterterrorism scholar Alex Schmid argued that, among the counterterrorism tools available to governments, inadequate attention had been given to what he termed “psychological—communicational—educational” initiatives. More could be done to advance public relations campaigns in the cause of counterterrorism, to provide fora for freedom of expression and to stress to political opponents a common aversion to being victimized by terrorist violence. “If one wants to oppose terrorism,” Schmid suggested, “the political communications of terrorists have to be opposed not by censorship but by a persuasive dialogue with various audiences including the supporters of terrorist movements. It is my impression that in countering terrorism, too much energy is placed on combating the violence, and not enough on countering the political propaganda of terrorists. It is the latter that gets terrorist movements new recruits and keeps their constituencies committed to their cause.”¹ Similarly, writing prior to 9/11 about the United States, Paul Pillar, a former intelligence officer and academic, contended that “[t]he attention to detail and nuance that is common to commercial marketing needs to be applied to what the United States tells the

outside world about terrorism, terrorists, and what the United States is doing about terrorism.”²

It is appropriate to revisit such claims today. If kinetic or “hard-power” approaches to counterterrorism predominated in the pre-9/11 period and, as has been amply argued, in the immediate post-9/11 years, the more recent trend has been to broaden the scope of counterterrorism approaches and to expand the use of “soft power” tools. The emergence of terrorism prevention strategies at the national and multilateral levels in recent years speaks directly to the imbalance noted above. It also reflects the evolution of the terrorist threat over this period. Whereas the dangers posed by al-Qaida as an organization appear to have receded, a greater diversity of threats, including homegrown extremists, command attention. Against this background, terrorism prevention policies have been pursued by a growing number of states under such different rubrics as counterradicalization, preventing violent extremism, and countering violent extremism (CVE). In the multilateral context, the terms “prevention of incitement,” per UN Security Council Resolution 1624, and “measures to address conditions conducive to the threat of terrorism,” per Pillar I of the UN General Assembly’s *United Nations Global Counter-Terrorism Strategy*, are also closely related. The Global Counterterrorism Forum

1. Alex P. Schmid, “Towards Joint Political Strategies for De-legitimising the Use of Terrorism” Proceedings of the International Conference on “Countering Terrorism Through Enhanced International Cooperation,” (Courmayeur Mont Blanc, Italy, 22–24 September 2000), pp. 261–262.

2. Paul R. Pillar, *Terrorism and U.S. Foreign Policy* (Washington, D.C.: Brookings Institution Press, 2001), p. 200.

(GCTF) refers directly to CVE and has established a thematic working group on that topic. It is anticipated that an international Center of Excellence on CVE will be established in the fall of 2012 under the auspices of the GCTF. Beyond the convergence of national counterterrorism policies around the idea of prevention, CVE is now part of the architecture of international counterterrorism cooperation.

What then is terrorism prevention? Although states pursue a range of operational measures to disrupt terrorist plots in the planning and execution phases, we use the term “terrorism prevention” in line with emerging practice to describe those measures designed to counter the ideas, narrative, or “message” advanced by extremists and complement operational preventive efforts. Terrorism prevention aims to “prevent non-radicalized populations from becoming radicalized. The objective is to create individual and communal resilience against cognitive and/or violent radicalization through a variety of non-coercive means.”³ As this definition implies, terrorism prevention is often derivative of the concept of radicalization, which itself is a relatively new addition to the counterterrorism vernacular. Radicalization has been widely recognized as a multistep process through which an individual or small group becomes imbued with extremist views and seeks to realize them through violence.⁴ In this context, prevention succeeds when the individual or group in question does not complete that process. More precisely, a distinction is sometimes made between “cognitive” and “violent,” or “behavioral” counterradicalization. The former refers to an ideological transformation stemming from a willingness to renounce a belief in violent extremism; the latter implies a disavowal of the use of violence or a change in behavior without necessarily rejecting extremist ideas.⁵

As this suggests, the key objectives of terrorism prevention are to persuade individuals that extremist ideas are wrong and dissuade them from acting on them. It follows that the audience for terrorism prevention policies is potentially vast, especially if these measures are directed to populations at home and abroad deemed to be vulnerable to providing ideological, financial, or operational support to violent extremists. For this reason, experience suggests that the range of noncoercive means that can comprise terrorism prevention efforts is quite broad. Consequently, states and multilateral organizations have pursued a wide range of activities to counter violent extremism.

In many cases, these measures have built on existing initiatives and policies (e.g., community engagement, development, conflict prevention, etc.) and are delivered by agencies that have not had a significant role in counterterrorism in the past or are unlikely to add the “counterterrorism” label to their current work. For this reason, some observers describe CVE as a policy “theme” rather than a policy per se, as it cuts across multiple areas of government action.⁶ Indeed, director of the London-based International Center for the Study of Radicalisation and Political Violence, Peter Neumann, notes that “[t]he range of activities that serve the aims of counterradicalization is potentially unlimited.” He offers a typology of indicative activities, comprising:

- messaging (through speeches, television programs, leaflets, social media, etc.);
- engagement and outreach (town halls, roundtables, advisory councils, etc.);
- capacity building (youth and women’s leadership initiatives, community development, community safety and protection programs, etc.); and
- education and training (of community leaders, public employees, law enforcement, etc.).⁷

3. Peter Neumann, “Preventing Violent Radicalization in America,” Bipartisan Policy Center, June 2011, p. 16, <http://bipartisanpolicy.org/sites/default/files/NSPG.pdf>.

4. Mitchell Silber and Arvin Bhatt, “Radicalization in the West: The Homegrown Threat,” New York City Police Department, 2007, http://www.nypdshield.org/public/SiteFiles/documents/NYPD_Report-Radicalization_in_the_West.pdf.

5. Tore Bjørge and John Horgan, eds., *Leaving Terrorism Behind* (Oxford: Routledge, 2009).

6. Neumann, “Preventing Violent Radicalization in America,” p. 18.

7. *Ibid.*

It follows that terrorism prevention initiatives, whether they are specifically labeled as such or not, are now underway in a variety of places, including prisons and detention centers, youth and sporting clubs, schools and universities, and churches, mosques, and other houses of worship. This brings counterterrorism officials into contact with a range of actors that were perhaps beyond the scope of their immediate work until relatively recently. In a policy domain that has been circumscribed and highly specialized in the past, the potential breadth of terrorism prevention underscores their novelty. In sum, terrorism prevention is a significant departure from the counterterrorism approaches of the pre-9/11 period. It mobilizes new implementing agencies from across government, engages a new range of nontraditional interlocutors outside of government, and extends counterterrorism to a series of policy domains (and physical spaces) that were not previously impacted by considerations of national security. There is much variation across states in their use of prevention measures and in their aims, budgets, and attempts to integrate prevention with other counterterrorism tools, but emerging policies and practices suggest that more states now view terrorism as requiring a whole-of-government-level response.⁸

The increased focus on prevention has brought with it greater recognition of the view that counterterrorism itself is a form of communication to vulnerable populations and the broader public.⁹ Writing about evaluating counterterrorism performance, Beatrice de Graaf has argued that the processes by which the threat of and response to terrorism are communicated to the public and the messaging these processes imply—their “performativity”—are closely linked to the levels of violence and radicalization. When performativity is low (e.g.,

counterterrorism efforts are conducted in a discreet manner without the language of a public “call to arms” and sensationalization of the threat), counterterrorism can have a more “rapidly neutralizing effect on radicalization.” Consequently, public perceptions of government performance in addressing the threat and securing citizens can provide one measure of effectiveness of counterterrorism and terrorism prevention efforts.¹⁰

A final definitional point concerns the distinction between terrorism prevention and the related concepts of deradicalization and disengagement. These terms pertain to individuals that have already become radicalized. Deradicalization is a cognitive transformation of a former terrorist away from terrorism while disengagement is a behavioral transformation, wherein an individual might continue to ascribe to extremist views.¹¹ Of course, the lines between deradicalization and disengagement on the one hand and terrorism prevention on the other may blur. For example, observing or participating in the disengagement process of a family member or friend may have the effect of dissuading others. Indeed, some deradicalization programs have been designed specifically to have a preventive effect on detainees’ families and friends and reduce the number of potential recruits for terrorist groups. Similarly, known terrorists (the subjects of disengagement and deradicalization measures) may be utilized directly or indirectly in the formulation and implementation of terrorism prevention policies. Therefore, although there is a conceptual distinction between counterterrorism and deradicalization, they may often be integrated in practice.

The development of prevention policies has given rise to a number of programmatic challenges for policymakers and practitioners.¹² These include, for example,

8. Lorenzo Vidino, “Countering Radicalization in America: Lessons From Europe,” *United States Institute of Peace Special Report*, no. 262 (November 2010), p. 2, http://www.usip.org/files/resources/SR262%20-%20Countering_Radicalization_in_America.pdf.

9. See Beatrice de Graaf, *Evaluating Counter-terrorism Performance: A Comparative Study* (New York: Routledge, 2011).

10. Beatrice de Graaf, “Why Communication and Performance Are Key in Countering Terrorism,” International Centre for Counter-Terrorism, 2011, <http://www.icct.nl/download/file/ICCT-de-Graaf-Communication-and-Performance-Key-to-CT-February.pdf>.

11. Tore Bjørgo and John Horgan, *Leaving Terrorism Behind* (Oxford: Routledge, 2009), pp. 1–10.

12. See Vidino, “Countering Radicalization in America”; Brian Fishman and Andrew Lebovich, “Countering Domestic Radicalization: Lessons for Intelligence Collection and Community Outreach,” New America Foundation, June 2011, http://www.newamerica.net/sites/newamerica.net/files/policydocs/Fishman_Lebovich_Domestic_Radicalization.pdf.

identifying community-level partners and framing interventions in such a way as to avoid the stigmatization of target audiences, which may include minority and faith-based groups. Similarly, policymakers have had to consider whether and how prevention policies address sensitive topics, such as religion, faith, and familial or civic duties. Target audiences are likely to be wary of the assertion of state authority in religious and family matters, particularly in states and societies where this is not the norm. More broadly, those elaborating counterradicalization policies must understand the evolving nature of radicalization and the factors that lead to extremism. In turn, they must define the aims of the intervention, i.e., whether to target both cognitive and behavioral radicalization.¹³ Some of these general programmatic challenges overlap with the specific challenges that arise in evaluating terrorism prevention policies. The latter has become an emerging concern for states in recent years. Evaluation can yield information about the costs, efficiency, and impact of such measures and is crucial in determining whether initiatives are proportionate to the threat and sustainable in the current financial and political climate. Put simply, having begun to invest in prevention programming, it is important for policymakers to ask whether these measures work and how experience can inform future practice.

13. Vidino, "Countering Radicalization in America: Lessons from Europe," pp. 4–6.

Evaluating Terrorism Prevention Programs: Challenges and Opportunities

ACROSS THE FIELDS OF PUBLIC POLICY AT THE national and international levels, program evaluation generally involves the identification of inputs, activities, outputs, outcomes, and impacts (box 1). As the use of those terms implies, evaluation is often conceived as a technical exercise, to ensure that resources are optimally and accountably deployed toward

the achievement of specified objectives. Although there may be multiple ways of approaching any particular evaluation, the importance of documenting the links between inputs to impacts, via activities, outputs, and outcomes, is a common denominator. Indeed, perhaps the critical link is that between outputs (what agencies do) to outcomes (what agencies hope to achieve).

BOX 1: KEY TERMS^a

Activities: Actions taken or work performed through which inputs are mobilized to produce specific outputs.

Evaluation: Systematic and objective assessment of an ongoing or completed project, program, or policy and its design, implementation, and results.

Impacts: Positive and negative, primary and secondary, long-term effects produced by an intervention, directly or indirectly, intended or unintended.

Inputs: Financial, human, and material resources used for a project, program, or policy.

Monitoring: Collection of data on specified indicators to provide

feedback on the consistency or discrepancy between planned and actual activities, the achievement of objectives, and the use of funds.

Outcomes: Actual or intended short-term and medium-term effects of an intervention's outputs, yielding the change in conditions that the intervention seeks to support.

In the field of counterterrorism generally, the measurement of outcomes and impacts is inherently difficult. Despite the massive investment of resources in this field, especially in the post-9/11 period, few states and multilateral organizations have elaborated robust and succinct methodologies to evaluate their counterterrorism measures per se. Likewise, the academic literature on the effectiveness of counterterrorism generally remains in its infancy. As participants in the Ottawa colloquium reflected, states have often had to begin from scratch in approaching the specific question of how to evaluate terrorism prevention measures. An initial challenge arises because the desired outcome is the absence of terrorism or violent extremism, so practitioners must measure a nonevent, an outcome that does not happen (radicalization). The challenge of “measuring the negative” warrants careful consideration at each stage in the evaluation process.

For these reasons, the Ottawa colloquium was designed to advance two objectives: to better understand existing approaches to the evaluation of counterterrorism efforts, in particular those geared at terrorism prevention, and to learn from practitioners in fields related to violence prevention or others where the challenges of measuring the negative have been confronted. This idea of “learning by analogy” is familiar to many states who, in their strategic documents on terrorism prevention,

acknowledge the relevance of existing programs on gang prevention, community engagement, immigrant integration, and so on. Evaluation efforts in these fields, especially where the broader objective is to dissuade vulnerable populations against the use of violence or to alter closely held beliefs or perceptions, yielded a number of valuable insights. In particular, experts who work on evaluation in domains that now may be considered part of the reach of counterterrorism, such as development aid and strategic communications, reflected on the methodologies and tools that have been and may be deployed to evaluate prevention initiatives.

Indeed, the commonality of challenges across the various fields explored during the colloquium underscored the value of opportunities to interact and to share practices and lessons learned among a wide range of practitioners and experts. For example, one presentation provided an overview of the evaluation process in the field of conflict transformation (fig. 1). This “wheel” framework disaggregates evaluation activities into three stages: design, implementation, and “evaluation,” i.e., assessment of data and of the evaluation process itself. A key element of this framework is the development and review of “theories of change” (box 1), which explicate or model the logic that connects inputs, activities, outputs, outcomes, and impacts. Clear specification of the “theory of change” entailed

Outputs: Tangible products that are directly attributable to initiatives, over which implementing agencies and stakeholders have the most control.

Performance measurement: Process of designing and implementing specific measures, i.e., metrics or indicators, of program results.

Theory of change, or program theory: Explication of the causal links from inputs, activities, and outputs through to outcomes and impacts; often depicted in logic models.

a. UN Development Programme Executive Board and the UN Population Fund, “The Evaluation Policy of UNDP,” DP/2011/3, 15 November 2010, paras. 9–10; Joy A. Frechtling, *Logic Modeling Methods in Program Evaluations* (San Francisco: John Wiley and Sons, 2007); 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, 2004), annex 6; James C. McDavid and Laura R.L. Hawthorn, *Program Evaluation and Performance Measurement: An Introduction to Practice* (Thousand Oaks, Calif.: SAGE Publications, 2006); Emil J. Posavac, *Program Evaluation: Methods and Case Studies* (Upper Saddle River, N.J.: Prentice Hall, 2011).

in terrorism prevention programming is important in deciding what specific measures, i.e., indicators and metrics, to utilize and what data to gather.

Figure 1: Evaluation Cycle

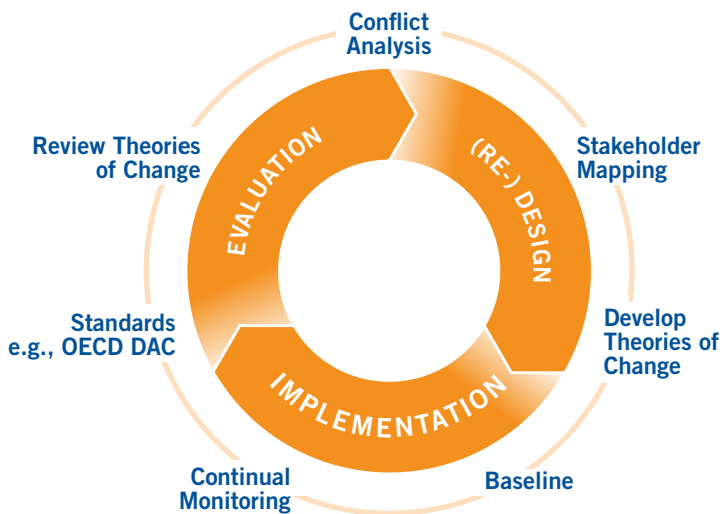


Diagram courtesy of Integrity Research and Consultancy

Figure 1 reflects the performance management cycle that is discussed across the evaluation literature.¹⁴ In the sections that follow, we summarize and elaborate on insights from the Ottawa colloquium under three headings reflected in the diagram: evaluation design, implementation, and analysis. We show that although practitioners in the field of terrorism prevention face significant challenges in evaluating programs, there are opportunities to learn from their initial experience and evaluation experts in related fields.

Evaluation Design

In considering evaluations of terrorism prevention programming, practitioners are faced with a series of threshold questions pertaining to evaluation design. Most broadly, evaluations may be used for different purposes. Scholars have identified “need” (for the program), “cost,” and “outcomes” as the fundamental drivers of evaluation activities.¹⁵ A clear understanding of evaluation objectives is needed to determine the type of evaluation to pursue. For example, in embarking on a new program, practitioners may opt to undertake an “evaluability assessment” to determine whether a program is ripe for evaluation. Some evaluations will aim to improve the delivery of a program (“formative” evaluations) while others will reflect on the impacts of completed programs (“summative” evaluations).¹⁶ If evaluation is to add value, objectives must be clearly stated up front.

Next, in light of the potential breadth of factors that impact prevention outcomes, practitioners must define the scope of an evaluation. As noted above, terrorism prevention policies are a relatively new addition to the counterterrorism tool kit, and we are at the early stages in accumulating knowledge about how prevention works in countering radicalization. Participants in Ottawa described the dilemma of attribution, i.e., drawing a line of causality between the desired outcomes that we observe (nonradicalization or nonviolence) and a specific prevention initiative. For this reason, practitioners have suggested basing evaluations on clearly elaborated theories of change that are themselves built on current understanding of the radicalization process and can be tested through program activities, a process also called logic modeling. Beyond verifying causation, logic modeling can aid practitioners in the critical task of identifying the specific measures, i.e., indicators

14. James C. McDavid and Laura R.L. Hawthorn, *Program Evaluation and Performance Measurement: An Introduction to Practice* (Thousand Oaks, Calif.: SAGE Publications, 2006), pp. 6–8.

15. Emil J. Posavac, *Program Evaluation: Methods and Case Studies* (Upper Saddle River, N.J.: Prentice Hall, 2011), ch. 1.

16. Kathryn E. Newcomer, Harry P. Hatry, and Joseph S. Wholey, “Planning and Designing Useful Evaluations,” in *Handbook of Practical Evaluation Design*, ed. Joseph S. Wholey, Harry P. Hatry, and Kathryn E. Newcomer (San Francisco: John Wiley and Sons, 2010), pp. 8–10.

and metrics, on which to gather data.¹⁷ In turn, practitioners must determine which tools and methods to use to gather data, whether using qualitative or quantitative tools, primary or secondary sources, and fieldwork or desk research, or mixed methods.¹⁸

In the design phase, practitioners must resolve some practical questions, for example, regarding the identification of an evaluator. Given the breadth of prevention measures, this may require the cooperation of agencies across government. Identifying an in-house evaluation team may hold appeal. Evaluators that are unfamiliar with the context or the program may elicit the wariness of program staff or cooperating partners, which may compromise the uptake of an evaluation outcome. At the same time, as one expert put it, “[y]ou can’t mark your own homework,” and it may be important to have external evaluators as they are in a position to provide an objective and independent assessment. In the foreground too is the question of resources. How much can and should be committed to evaluation in light of the anticipated utility of evaluation results? Tracing the causal impacts of programs can be a complex and costly task, and evaluations should be designed to be feasible as well as useful.

Reflected in the presentations in Ottawa, there is a growing body of evidence that states have considered these questions and begun to pursue measures to identify successful counterterrorism and prevention actions. The means and methodologies for doing so have been tied in large part to the desired objective of the evaluation. Presenters identified cost effectiveness, impact on the community, and proportionality, i.e., whether the means of suppressing radicalization are properly aligned with current threat assessments, as key objectives of evaluations undertaken to date. In relating their experiences in defining the scope of evaluations, practitioners

highlighted the importance of deciding on a few key points prior to the evaluation. These include a need to decide whether the evaluation would take a strategic or operational approach, consider structural or proximate factors, and examine a single project or a broader set of policies.

Participants noted that a number of policy areas and activities may contribute to the overall objective of preventing radicalization and terrorism. Nevertheless, evaluators must decide which programs are to be evaluated and how to account for the impact of those programs outside the assessment. This issue is particularly relevant where the contributions of multiple activities and agencies are being considered (the question poses less of a challenge where a specific program or project is being evaluated). As one speaker noted, an overall improvement in public perceptions of security, the absence of attacks, and decreased recruitment into violent extremist groups all combine to produce a favorable assessment of counterterrorism and CVE efforts. A number of policy areas, however, including education, law enforcement, development, and diplomacy may have contributed to these efforts. Qualitative data and surveys that capture public perceptions about security, community relations, and support for violent extremist groups and ideologies can contribute to such assessments. Relatedly, it is important at the outset to have realistic expectations regarding the impact of an intervention. In areas where there are few prospects for change, small shifts may constitute large successes.

Regarding specific measures, participants surveyed possible metrics and indicators of counterterrorism “success.” There has not been another attack of the scale and complexity of 9/11, although tragic attacks in Bali, Madrid, London, and elsewhere have testified to the continuing interest of terrorists to propagate

17. On the benefits of using logic models and “theories of change,” see John A. McLaughlin and Gretchen B. Jordan, “Using Logic Models,” in *Handbook of Practical Evaluation Design*, ed. Joseph S. Wholey, Harry P. Hatry, and Kathryn E. Newcomer (San Francisco: John Wiley and Sons, 2010), ch. 3; Posavac, *Program Evaluation*, ch. 3; Joy A. Frechtling, *Logic Modeling Methods in Program Evaluations* (San Francisco: John Wiley and Sons, 2007), ch. 2.

18. For an overview, see Anthony Ellis et al., “Monitoring and Evaluation Tools for Counterterrorism Program Effectiveness,” *Center on Global Counterterrorism Cooperation Policy Brief*, September 2011, http://www.globalct.org/images/content/pdf/policybriefs/integrity_policybrief_1111.pdf.

mass-casualty attacks. One way to measure effectiveness and identify success, it was suggested, is to consider what a lack of success would look like. It is likely that terrorist attacks would be more frequent and deadly that fear would become a permanent element of public life and lead to the securitization of more functions and that perpetrators would be able to exploit capacity gaps in law enforcement and enjoy impunity. Noting the tensions between a reduction in terrorism and a general increase of fear and anxiety in the public domain, one speaker suggested that overall there was a mixed picture of success.

Depending on the metrics and indicators identified, a wide range of methods and tools should be considered to gather the data. These may include surveys, focus groups, content analysis (of media, websites, and such), and interviews, among others. Throughout this process, practitioners should consider the question of how the data gathered will be stored for purposes of future analysis. Database development is an important consideration in determining what and how to measure.

In evaluating terrorism prevention programs, a particular challenge pertains to states confronted with the threat of terrorism from multiple sources, including groups and individuals with separatist, religious, or ideological objectives. Consequently, each is likely to require a different set of preventive measures and tailored programming, each with specifically related objectives and, therefore, indicators of effectiveness. The same evaluation tools may not always be suitable to the variety of activities this challenge requires. Moreover, national priorities are shaped by an ever-evolving assessment of strategic security threats. Responsive programs need to be adaptable, often making medium- and long-term evaluations more problematic.

Terrorism prevention programs developed against a set of well-defined theories of change become more evaluable. Although it is difficult to define the factors that lead to extremism, developing a suitable theory of change has helped in establishing benchmarks and

indicators against which activities might be evaluated. Subsequent program modifications and redesign could then reflect the lessons learned regarding the validity of the theory of change put forth at the outset. Yet, developing and refining theories of change takes time, and where evaluations have to take place in conflict zones or under great insecurity, attempts to compress timelines can affect the utility of evaluation.

Drawing on these themes, a principal insight gained from the Ottawa colloquium is that state experiences in evaluating terrorism prevention programs can be categorized as multidimensional, vertical, or horizontal.

Multidimensional evaluations utilize a framework that considers multiple levels of evaluation. For example, one state's CVE program evaluation framework has a five-level hierarchical model:

1. Line of Sight Evaluation: to evaluate if projects are aligned with CVE program outcomes and capabilities
2. Project Effectiveness Evaluation: to evaluate if projects meet their stated objectives
3. Outcomes Evaluation: to evaluate if projects contribute to program outcomes
4. Program Benefit Evaluation: to evaluate if program outcomes result in desired benefits
5. Program Impact and Evaluation: to evaluate if program benefits help realize the vision

Utilizing a scale from 1 to 5, the evaluations range from those undertaken in the short term to more medium- and long-term assessments.

Another form of multidimensional evaluation takes place, for example, if a government agency is asked to provide assessments and policy guidance on a wide variety of programs. In one state, for example, one entity is tasked with providing strategic assessments of a wide range of programs and provides policy guidance based on the evaluations. To fulfill this mission, they have used a combination of "grounded theory" and "framework analysis."¹⁹ Grounded theory reflects an iterative

data collection process in which inputs are drawn from a variety of data collection methods. Coding and categorization takes place during data collection and an argument, or theoretical paradigm, results directly from the patterns embedded in the data.

Framework analysis is designed to process existing data in five stages: (1) familiarization with the data; (2) identification of a thematic framework for understanding the data, i.e., development of a coding system that emerges from the data itself; (3) indexing and charting the data; (4) mapping the range and nature of the data; and (5) interpreting the data to provide a coherent explanatory model. These two approaches shape the analysis of data through literature reviews, structured or semistructured interviews, and focus groups, which inform resultant strategic impact assessments and strategic assessments. The former refers to assessing progress against specific government action plans or strategies, while the latter may address an open-ended question posed by a senior policymaker, follow-up to a strategic impact assessment, or an increasingly important area of counterterrorism policy that may have been overlooked by policymakers.

Vertical evaluations assess specific programs from inception to outcome. Through this approach, states develop programs to focus on those elements believed to pose the most imminent threats to national security and stability. The monitoring and evaluation of such programs requires a careful analysis of the desired objective, the costs, and institutional capacities for implementation. Additionally, wanted and unwanted side effects require careful consideration, as programs may generate community tensions, further suspicion of government intentions, and the securitization of social programs that may become compromised as a result. For example, one government program was

revised when it was felt that it promoted objectives that were no longer in line with national priorities and that funding for supporting projects was generating tensions in the target and other communities.

Evaluations of deradicalization programs would come under this category. There has been much debate about the need and value for universal indicators for success of such programs. Academics have underscored the need for comparable data in order to assess success. Some practitioners have questioned whether universal indicators are necessary, arguing that success for them is determined by a set of internal or national factors.²⁰ As indicated earlier, however, although many of these programs address detainees and hence where counter-radicalization presumably has failed, they have a preventive element. Emphasizing that aspect in the evaluation of the programs would help policymakers and program designers better understand and measure the effect of such efforts on preventing the radicalization of family members and others.

Horizontal evaluations examine the efforts undertaken by multiple agencies and entities to support a specific government action plan or strategy. The United States, for example, for the first time outlined a national strategy for addressing domestic violent extremism, *Empowering Local Partners to Prevent Violent Extremism in the United States*, and a strategic implementation plan that outlines a whole-of-government approach to this common objective. A first step, as outlined by the Strategic Implementation Plan, will be the development of key benchmarks to guide initial assessments. As the strategic plan matures, evaluations will shift toward impact assessments. Individual departments and agencies will be responsible for assessing their own programs in pursuit of the common objectives, in coordination with an assessment working group.

19. National Counterterrorism Center official, discussions and e-mail correspondence with authors, Washington D.C. and New York, February–April 2012.

20. Naureen Chowdhury Fink and Hamed El-Said, “Transforming Terrorists: Examining International Efforts to Address Violent Extremism,” International Peace Institute, May 2011, http://www.ipacademy.org/media/pdf/publications/2011_05_trans_terr_final.pdf; John Horgan and Kurt Braddock, “Rehabilitating the Terrorists? Challenges in Assessing the Effectiveness of De-radicalization Programs,” *Terrorism and Political Violence* 22, no. 2 (2010): 267–291.

Another example of a horizontal process is in the Netherlands, where an evaluation of the necessity, proportionality, and efficacy of a decade of counterterrorism measures following 9/11 was evaluated. This assessment was undertaken in response to a parliamentary request and the report of the Suyver Committee, which generated 10 key questions to provide the framework for evaluations:

1. To what extent is the new policy based on threats, practical experience, and international agreements?
2. To what extent are measures applied in counterterrorism practice, and do these measures offer value added in practical terms?
3. What are the opinions of the courts, independent advisory bodies, and other reviewers concerning the application of measures, and are these opinions taken into account when creating new policies?
4. How is the decision to apply one measure, rather than another, taken?
5. If the chosen measure had not existed, would any of the other existing measures have provided a remedy?
6. To what extent is the practical application of the measure in keeping with the objectives of this policy?
7. To what extent is the application of the measure dependent on the use of other measures, and what is the effect of this?
8. To what extent and in what way does the safeguarding of legal rights play a role in the application of the measure?
9. To what extent and in what way is unity of command of relevance to the effective application of the measure?
10. To what extent and in what way does cooperation play a role in the application of these measures?

A combination of approaches was used to provide a practical evaluation in a relatively short period. A generic approach led to the development of a literature review and supplementary interviews; specific measures

were then examined in more detail. The case study approach involved interactive thematic sessions with practitioners to evaluate particular cases, and a case law approach examined the relevant legislative measures. These four approaches were applied to studying the cause, application, assessment, and subsequent amendment of national counterterrorism measures.²¹

Implementing Evaluations

As with program design, the process of implementing evaluations raised a number of common challenges and approaches. For example, once a decision has been taken regarding what should be measured and how, a series of methodological and practical challenges arise regarding the actual collection of data. An initial point concerns the question of when to begin gathering data. The collection of baseline data prior to the intervention allows for a before-and-after comparison of the impact of prevention measures. Such baselining is the most robust way of determining the causal effects of prevention programming. Some states have used polling as a method to gather baseline data in evaluating terrorism prevention programs. As an alternative or in addition to baselining, the causal effects of programs can be tracked using control groups where possible. To measure the impact of development interventions, for example, one agency compared a community in which a number of programs were being implemented to another community of similar size and demographics in which intervention was minimal. This approach addressed to some extent the challenge of attribution, linking outputs and outcomes.

A second cluster of concerns in implementing evaluations pertains to stakeholder engagement, a topic that has attracted increasing attention in the evaluation

21. Government of the Netherlands, "Counterterrorism Measures in the Netherlands in the First Decade of the 21st Century: On the Cause, Application, Assessment, and Amendment of Counterterrorism Measures in the Netherlands 2001–2010," January 2011, http://www.coe.int/t/dlapiil/codexter/Source/Working_Documents/2011/CT_Measures_Netherlands.pdf.

literature in recent years.²² Program staff, for example, may be resistant to evaluation for several reasons. Evaluation processes may be seen as taking them away from project implementation, as threatening to their jobs or income, or as an external imposition that does not align with local needs and priorities. Where program staff members were unable to participate in the development of theories of change, benchmark indicators, or program design, they were often wary of being evaluated against criteria that were seen as beyond their control. Government agencies shared similar concerns, noting, for example, that events and initiatives beyond their remit could affect their performance. These issues underscored the value of working with evaluators who are familiar with the context, the programs, and some of the key stakeholders.

Evaluation experts that have experience in working in conflict or postconflict environments emphasized that evaluation processes may seem alien to the proximate needs and priorities of officials and civilian populations. Local ownership can be a key factor in determining the success of an evaluation and ensure that the outcome is taken into consideration in the development of future programs and policies. Additionally, ensuring that local actors are involved in the evaluation process can provide opportunities for institutional capacity-building and the development of evaluable programs that can be monitored on a more sustained basis by domestic practitioners. The importance of being context sensitive and conflict sensitive, where relevant, underscored the importance of working with local partners to develop ownership and ensure that evaluators are well educated about project or program, as well as the environment that is being assessed. This engagement with local actors was seen as an integral means of ensuring the necessary buy-in of stakeholders during the evaluation and the take-up of its outcomes

and recommendations, including in any subsequent program redesign.²³

More generally, the role of the evaluator warrants careful consideration in the implementation phase. Participants recalled that the process of gathering data poses significant challenges in some environments. For example, in conflict zones there may be constraints regarding access to and security for evaluation staff. This can lead to the compression of timelines and, in turn, imperfect data. In the field, whether at home or abroad, evaluators may themselves be working with vulnerable populations or within communities that are sensitive to initiatives advanced under the label of counterterrorism. The presence of evaluators can itself be viewed as a form of strategic communications in this regard. Practitioners should take steps to minimize the risk that data will be biased as a result of wariness among research subjects and to ensure that the broader policy objectives of terrorism prevention will not be undermined by the work of evaluators.

For these reasons, there is a need for pragmatism in implementing evaluations, including flexibility in the process of gathering data. Although evaluation processes are often associated with quantitative data, practitioners are increasingly placing value on qualitative data and, in some cases, anecdotal evidence. For example, the effect of an intervention on online activity may be difficult to quantify, but experts in Ottawa noted that it was possible to perceive changes in the discussions and nuances in the discourse, which might indicate the impact of an intervention, pointing out that “not everything that counts can be counted.” The challenge of effecting behavioral and ideological changes in complex and dynamic contexts has led to the use of mixed methods. For example, a variety of tools, including opinion polling, focus groups, media monitoring, cul-

22. For example, see Posavac, *Program Evaluation*, ch. 3; John M. Bryson and Michael Q. Patton, “Analyzing and Engaging Stakeholders,” in *Handbook of Practical Evaluation Design*, ed. Joseph S. Wholey, Harry P. Hatry, and Kathryn E. Newcomer (San Francisco: John Wiley and Sons, 2010), ch. 2.

23. Many of these concerns are also reflected in designing evaluations in the related field of security system reform. See Simon Rynn and Duncan Hiscox, “Evaluating for Security and Justice: Challenges and Opportunities for Improved Monitoring and Evaluation of Security System Reform Programmes,” Saferworld, December 2009, <http://www.saferworld.org.uk/downloads/pubdocs/Evaluating%20for%20security%20and%20justice.pdf>.

tural insights, and desktop research, could be combined to deliver a multidimensional assessment of impact.

The need for pragmatism was further underscored in circumstances where resources are likely to be a concern in implementing evaluations, just as they are in evaluation design. Experience suggests that this may lead evaluators to adjust or alter their approach or to utilize more cost-effective tools and methodologies.

A further point concerns multilateral organizations, many of which have drawn on evaluation to assess their work, and have confronted a unique set of challenges. As indicated above, terrorism prevention largely falls within the domain of Pillar I of the Strategy. The efforts of member states and UN agencies are supported and coordinated by the UN Counter-Terrorism Implementation Task Force (CTITF). Evaluating preventive programs in this case is not feasible as it would mean conducting a multidimensional evaluation of nearly 200 states. One option might be to collate existing indicators of relevance to Pillar I, such as the Millennium Development Goals, and evaluate states' performance on those topics accordingly. One example might be the set of indicators developed to assess implementation of Security Council Resolution 1325 on women, peace, and security.²⁴ Alternatively, the CTITF might consider asking states to report on Strategy implementation, as it does on a biennial basis, against a set of questions or indicators that will allow for some broad analysis about the state of Strategy implementation and impact. Another means of addressing the complex challenge of evaluating member states' counterterrorism capacities or programs has been the redesign of the preliminary implementation assessment by the UN Security Council Counter-Terrorism Committee Executive Directorate, tasked with monitoring global implementation of Security Council Resolution 1373, which, among other things, aims for the prevention of further terrorist attacks.

Analysis

A final series of challenges emerges in analyzing the data that has been gathered: What does the evidence say about the impact of prevention initiatives? How should these findings be presented to decision-makers?²⁵ Consideration of policies and practices across a range of public and private sector activity, including advertising and public relations, indicated that it is a struggle to understand the impact of an intervention. One presenter recalled the famous remark of Lord Leverhulme, founder of Unilever, who once noted that “[h]alf of my advertising money is wasted. The problem is, I don't know which half.”

Chief among the challenges that practitioners encounter in the analysis phase is causality attribution, which involves assessing data in light of the theories of change specified at the outset. Ideally, analysis will show that the specific prevention measures pursued led to the projected outcome, but the picture is likely to be more complicated. Specific prevention initiatives may yield unintended consequences. The response of terrorists to prevention measures is especially pertinent. Terrorism prevention entails an ideational contest with extremists, who are likely to frame these interventions in such a way as to reinforce their existing narratives toward the goal of attracting followers to their cause. We can anticipate that a tiny core of incorrigible extremists will endeavor to use prevention policies against the states pursuing them. Although governments can control the extent to which they stay on message in reinforcing the case against violent extremism (a challenge in itself), there are no guarantees that prevention policies will be interpreted in the way they are intended. There is some evidence of this from recent efforts to counter extremist messages on the Internet, that is, extremists have made efforts to blunt the counter-narrative advanced by some governments.

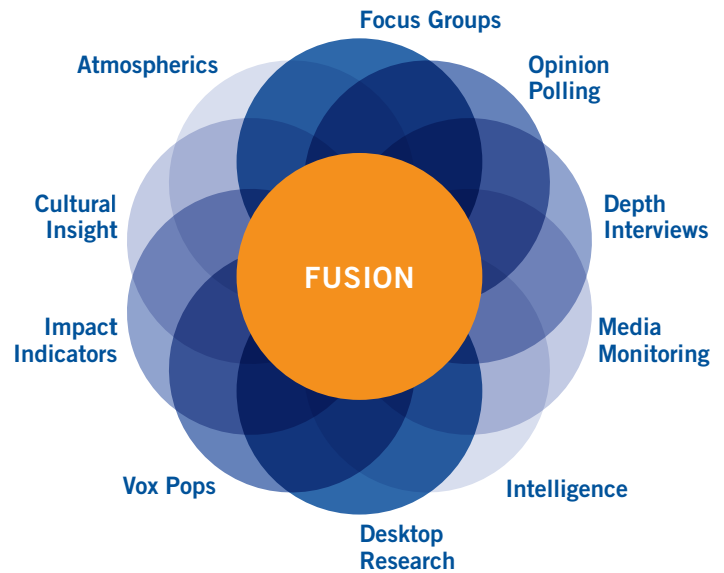
24. Security Council Report, “Cross-Cutting Report on Women, Peace and Security,” 2012 No. 1, 27 January 2012, p. 13.

25. On this point, the possible political implications of evaluation reporting were acknowledged by participants. For further discussion, see Posavac, *Program Evaluation*, ch. 13; Joseph S. Wholey, “Use of Evaluation in Government: The Politics of Evaluation,” in *Handbook of Practical Evaluation Design*, ed. Joseph S. Wholey, Harry P. Hatry, and Kathryn E. Newcomer (San Francisco: John Wiley and Sons, 2010), ch. 27.

It follows that, in interpreting the costs and benefits of prevention initiatives, analysts must balance the intended and unintended consequences of interventions. What standard should be set for success, i.e., at what point should we conclude that the impact of prevention policies is positive? The case for pragmatism can be made especially strongly here. The challenges that practitioners face at the programmatic level, as well as in evaluating prevention measures, suggest that a relative standard of progress over time is perhaps most appropriate. In making such an assessment, a willingness to reconcile the conceptual and methodological elegance of evaluation in theory with the real world of evaluation in practice is particularly helpful.²⁶ As a result, the dilemmas of measuring the negative may require a departure from ideal evaluation practices. As one practitioner noted, “The world is just too complicated, and there are too many competing ideas that can influence outlook and behavior. But if we look at the weight of available data and fuse multiple impact indicators, we can have confidence that the ‘direction of travel’ is heading in the right direction.” Another presenter introduced the ideas of “fusion” (fig. 2) and “bricolage” to describe the process of using diverse types evidence to make informed, professional judgments about the effects of specific interventions.

Our discussions and research suggest that evaluation is not simply a deductive exercise. Rather, evaluation is a learning process in itself, and robust analysis of data provides the opportunity to close the loop, inducing lessons learned and feeding them back into future efforts to conceptualize, operationalize, and implement evaluations. Lessons learned are particularly important given variability in processes of radicalization over time. Indeed, states will likely require a willingness to experiment and the flexibility to adjust to these challenges.

Figure 2: Multidimensional Indicators



Source: Bell Pottinger Public Advocacy, “Measuring the Negative: Insights From Public Relations,” February 2012.

26. McDavid and Hawthorn, *Program Evaluation and Performance Measurement*, pp. 407–410.

Further Research and Follow-Up

THE MOVE TO TERRORISM PREVENTION IS A TIMELY and important development in counterterrorism policy. This policy report provides some initial reflections on the question of performance measurement in this new and expanding field. By their own account, states have been “learning by doing,” but further work is needed. Among the outcomes of the colloquium is a better sense of what kinds of further research and analysis are needed to inform the development of necessary policies and practices among stakeholders. The following recommendations offer some initial ideas not only for further developing evaluation tools and responding to some of the challenges outlined above, but also for establishing platforms for greater interaction and cooperation among key stakeholders to help shorten learning cycles and develop pools of comparable data, where appropriate.

Mapping exercise. Although this report has offered an initial survey of the approaches taken by some governments and practitioners toward monitoring and evaluating terrorism prevention programs, a more comprehensive mapping exercise could survey and record existing evaluation practices of a broad range of states and practitioners and develop a set of models which may be applied in other areas in whole or in part, as suitable to the context.

Database. Over time and with continued investment, the initial mapping exercise could evolve into a database recording experiences, common methodologies, and indicators in order to shorten learning cycles and eliminate the need to develop evaluation tools from scratch. The ability to draw on data gathered by others may be particularly useful for policymakers from different states advancing prevention initiatives in the same foreign locations. Similarly, the findings of other states or organizations may be relevant to partners seeking to establish the effectiveness of their own programming. In related fields where evaluation efforts have been addressed in more detail, such a database could serve as a portal to further information and relevant publications or web-based tools.

Training. The interest in exchanging information and good practices suggested the need for greater training, and an investment in providing capacities and guidance for evaluating counterterrorism and CVE efforts was highlighted by a number of speakers. These would eliminate the need to develop new methods each time an evaluation was needed and contribute to some uniformity across evaluation practices, thereby facilitating comparative analysis, and, when including local partners, can contribute to capacity-building and institutional strengthening efforts.

Platforms for information sharing. A series of workshops could help deepen and broaden discussions among governments, experts, and members of relevant international organizations, such as the United Nations and the GCTF, to share data, lessons learned, and good practices regarding evaluation practices. Topics addressed by these workshops could include identification of communities vulnerable to violent extremist ideologies and recruitment, the key target audiences for CVE programs; development and testing of theories of change (models and experiences); practices relating to measuring preventive interventions in violence-related fields; and practices relating to measuring the negative in other fields.

Developing evaluation tools for the Internet.

Isolating and quantifying the impact of preventive programming, or measuring the negative, may be difficult, but the Internet provides a platform through which impacts may be more easily captured through statistics or website content. A content analysis and study of extremist websites could track the discourse before and after relevant CVE measures and offer some initial indicators about which measures had been most effective and how to develop evaluation tools to capture that kind of information. The study also can help indicate which entry points in the process of radicalization have proved most fertile for CVE measures.

A tool kit for policymakers and evaluators.

As interest in performance measurement in the field of terrorism prevention grows, there will be a need for training and tools to develop capacity and provide guidance in evaluating counterterrorism and CVE efforts. Beyond collating existing national practices, an evaluation tool kit for terrorism prevention initiatives would provide evaluators with a menu of options concerning metrics, methodologies, theories of change, and more that could be applied in different contexts. This would help

shorten learning cycles, eliminate the need to develop new methods each time an evaluation was needed, and contribute to some uniformity across evaluation practices, thereby facilitating comparative analysis.

In sum, as the threat of violent extremism evolves, so counterterrorism and CVE efforts must evolve with it. A more robust understanding of the effectiveness of particular tools, such as terrorism prevention, is essential in delivering a counterterrorism policy that is balanced and effective.

Suggested Readings

1. Naureen Chowdhury Fink and Hamed El-Said, "Transforming Terrorists: Examining International Efforts to Address Violent Extremism," International Peace Institute, May 2011, http://www.ipacademy.org/media/pdf/publications/2011_05_trans_terr_final.pdf.
2. Beatrice de Graaf, *Evaluating Counter-terrorism Performance: A Comparative Study* (New York: Routledge, 2011).
3. Anthony Ellis et al., "Monitoring and Evaluation Tools for Counterterrorism Program Effectiveness," *Center on Global Counterterrorism Cooperation Policy Brief*, September 2011, http://www.globalct.org/images/content/pdf/policybriefs/integrity_policybrief_1111.pdf.
4. John Horgan and Kurt Braddock, "Rehabilitating the Terrorists? Challenges in Assessing the Effectiveness of Deradicalization Programs," *Terrorism and Political Violence* 22, no. 2 (2010).
5. Cynthia Lum, Leslie W. Kennedy, and Alison J. Sherley, "The Effectiveness of Counter-Terrorism Strategies," *Campbell Systematic Review* 2006, no. 2 (January 2006), <http://www.campbellcollaboration.org/lib/download/53/>.
6. Government of the Netherlands, "Counterterrorism Measures in the Netherlands in the First Decade of the 21st Century: On the Cause, Application, Assessment, and Amendment of Counterterrorism Measures in the Netherlands 2001–2010," January 2011, http://www.coe.int/t/dlapil/codexter/Source/Working_Documents/2011/CT_Measures_Netherlands.pdf.
7. Peter Neumann, "Preventing Violent Radicalization in America," Bipartisan Policy Center, June 2011, <http://bipartisanpolicy.org/sites/default/files/NSPG.pdf>.
8. Raphael Perl, "Combating Terrorism: The Challenge of Measuring Effectiveness," *CRS Report for Congress*, RL33160, March 2007.
9. Emil J. Posavac, *Program Evaluation: Methods and Case Studies* (Upper Saddle River, N.J.: Prentice Hall, 2011).
10. Jason Rineheart, "Counterterrorism and Counterinsurgency," *Perspectives on Terrorism* 4, no. 5 (November 2010).
11. Simon Rynn and Duncan Hiscox, "Evaluating for Security and Justice: Challenges and Opportunities for Improved Monitoring and Evaluation of Security System Reform Programmes," Saferworld, December 2009, <http://www.saferworld.org.uk/downloads/pubdocs/Evaluating%20for%20security%20and%20justice.pdf>.
12. Alexander Spencer, "The Problems of Evaluating Counter-Terrorism," *UNISCI Discussion Papers*, no. 12 (October 2006).
13. The White House, "Empowering Local Partners to Prevent Violent Extremism in the United States," August 2011, http://www.whitehouse.gov/sites/default/files/empowering_local_partners.pdf.
14. Joseph S. Wholey, Harry P. Hatry, and Kathryn E. Newcomer, eds., *Handbook of Practical Evaluation* (San Francisco: John Wiley and Sons, 2010).



Center on Global
Counterterrorism
Cooperation