



NATIONAL INSTITUTE OF JUSTICE
**EVALUATING AND ASSESSING
TERRORISM PREVENTION PROGRAMS**

What Research Sponsored by the
National Institute of Justice Tells Us

Kateira Aryaeinejad and Thomas Leo Scherer, Ph.D.

APRIL 2024

U.S. Department of Justice
Office of Justice Programs
810 Seventh St. N.W.
Washington, DC 20531

Nancy La Vigne, Ph.D.

Director, National Institute of Justice

This and other publications and products of the National Institute of Justice can be found at:

National Institute of Justice

Advancing Justice Through Science

NIJ.ojp.gov

Office of Justice Programs

Building Solutions • Supporting Communities • Advancing Justice

OJP.gov

The National Institute of Justice is the research, development, and evaluation agency of the U.S. Department of Justice. NIJ's mission is to advance scientific research, development, and evaluation to enhance the administration of justice and public safety.

The National Institute of Justice is a component of the Office of Justice Programs, which also includes the Bureau of Justice Assistance; the Bureau of Justice Statistics; the Office for Victims of Crime; the Office of Juvenile Justice and Delinquency Prevention; and the Office of Sex Offender Sentencing, Monitoring, Apprehending, Registering, and Tracking.

Opinions or points of view expressed in this document represent a consensus of the authors and do not necessarily represent the official position, policies, terminology, or posture of the U.S. Department of Justice on domestic violent extremism. The content is not intended to create, does not create, and may not be relied upon to create any rights, substantive or procedural, enforceable at law by any party in any matter civil or criminal.

Executive Summary

In 2012, the National Institute of Justice's (NIJ) Domestic Radicalization to Terrorism program began funding research on issues related to domestic radicalization and terrorism in the United States. A component of this portfolio is focused specifically on evaluations and assessments of terrorism prevention programs implemented across different populations. This report synthesizes these efforts, reviewing shared and singular lessons from NIJ-sponsored evaluations of terrorism prevention and terrorism prevention-relevant programming and focusing on both evaluation findings and processes.

The projects reviewed for this report vary notably in their timelines, methods, programmatic focuses, and audiences. As such, the lessons drawn from them are not necessarily scientifically comparable or generalizable. However, important lessons emerge from this research, underscoring shared insights from NIJ-sponsored evaluative efforts and illuminating important areas for consideration in future programmatic and evaluative endeavors.

First, findings from NIJ-sponsored evaluations emphasize the importance of gaining community buy-in and assuring program relevance for participants before implementing terrorism prevention or terrorism prevention-relevant programs. Although this may vary based on the intended program audience (e.g., youth, communities, law enforcement), the findings overall suggest that to ensure program fidelity, utility, relevance, and buy-in, programs should be developed in close consultation and cooperation with the audiences for which they are constructed. Doing so can also help alleviate concerns regarding the perceived stigmatization or targeting of specific communities in which terrorism prevention efforts take place and help address concerns about and issues in associating specific communities or demographics with national security threats.

Second, programs for terrorism prevention and preventing/countering violent extremism can yield benefits beyond meeting terrorism prevention goals. They can also be complementary to and even carried out within public health or community resilience initiatives. The utility of incorporating terrorism prevention into these frameworks may vary based on the intended program audience and outputs. Careful attention should be paid in framing program activities and language when they are implemented via these larger frameworks. Indeed, if violent extremism awareness and prevention efforts are not appropriately integrated into a public health or community resilience model, participants may be confused as to the purpose, goals, and nature of the activities in which they are participating, to the detriment of short- and long-term programmatic goals and awareness-raising activities.

Third, evaluative efforts in general face programmatic and methodological challenges that limit their ability to assess the impact of a program and the generalizability and veracity of their findings. NIJ-sponsored evaluations went to commendable lengths to address issues associated with data availability, generalizability, validity, and determinations of impact; however, certain challenges — including participant attrition rates and lack of available data — limited their ability to do so. These challenges offer important lessons for future evaluations of terrorism prevention programs.

These evaluation findings suggest that several activities should be considered in addressing challenges that may arise during the evaluation process. These activities could include implementing formative evaluations or evaluability assessments prior to program implementation and evaluation, implementing strategies to limit participant attrition in the interest of maintaining appropriate sample sizes, and ensuring adequate resourcing for continued engagement in terrorism prevention initiatives. In addition, NIJ-sponsored evaluation findings suggest that program evaluators should consider incorporating control or comparison groups to assess programmatic impact on individual attitudes, behaviors, and outcomes, along with comparative analysis focused on assessing the outcomes and impact of programs replicated in different communities and settings to increase the generalizability, validity, and utility of evaluation findings. More information about NIJ-sponsored efforts to incorporate these elements is detailed in this report.

Looking forward, findings from this review suggest important gaps and considerations that should be addressed by and incorporated in future programmatic and evaluation decisions. These include evaluating the long-term impact of programs through follow-up assessments, incorporating additional indicators for measuring and assessing potential behavioral change and knowledge retention, and assessing the comparability of programs' impact on different forms of terrorism across different ideological spectrums.

Finally, further consideration should be given to determining the appropriate approach, strategy, and goals for terrorism prevention activities based on the audiences they are geared toward. Based on the evaluation findings, in some cases it may be beneficial to explore and test efforts aimed at incorporating terrorism prevention activities and programs within a broader portfolio of public health and violence reduction-focused efforts. Future programmatic, research, and policy-oriented activities might consider further exploration of the benefits and potential issues associated with doing so in addressing radicalization to terrorism within the United States.

Introduction

In 2012, the National Institute of Justice's (NIJ) Domestic Radicalization to Terrorism program began funding research on issues related to domestic radicalization and terrorism in the United States.¹ As these projects have come to completion, NIJ has organized the results into synthesis reports to highlight the most salient findings and useful lessons. Past reports have focused on investigating domestic radicalization processes and the potential risk factors and indicators associated with them.² This report examines lessons learned to date from NIJ-funded evaluations and assessments of programs focused on preventing radicalization processes and terrorism from occurring within the United States.³

Program evaluations are crucial to ensuring the efficacy, suitability, implementation fidelity, and impact of terrorism prevention efforts. Not only do they play an important role in measuring the program's impact and uncovering any unintended outcomes, but evaluations also yield insights valuable to the adaptation and alteration of programs to better achieve terrorism prevention goals and serve the needs of communities impacted by them. This report focuses on both lessons for terrorism prevention programs and lessons for future assessment and evaluation processes.

The report begins with a discussion of the importance of evaluative efforts in understanding and deriving lessons from terrorism prevention programs for policy and practice, as well as the inherent issues and challenges facing evaluations in this field. Following this discussion, the report provides an overview of NIJ-sponsored terrorism prevention evaluation research, spotlighting three NIJ-funded evaluations of programs implemented for community, youth, and law enforcement audiences between 2012 and 2019.⁴ Finally, collective lessons shared across NIJ-funded evaluations related to both programmatic findings and the evaluation process are presented before a discussion of the relevance of those lessons and their limitations, remaining gaps, and recommendations for future policy and practice.

Importance and Use of Evaluations

Evaluative efforts consist of a series of measures and indicators designed to assess the efficacy, feasibility, impact, or implementation of an activity or program in achieving set goals over a period of time.⁵ Evaluations differ in terms of their focuses and take varied forms depending on their intent, the availability of data, and the stage of programming at which they take place. These forms can range from evaluability assessments and formative evaluations focused on determining whether a program is ripe for evaluation,⁶ to process evaluations focused on programmatic implementation,⁷ to outcome and impact evaluations focused on measuring the effect of a program on its participants.⁸

Every type of evaluation contributes important insights and knowledge necessary to understanding the feasibility and real or potential impact of programmatic efforts, identifying areas for program improvement or modification, and assessing the extent to which the continuation of a program will result in the achievement of its stated goals. Understanding the suitability and efficacy of programming is of particular importance in terrorism prevention. Not only are such measures necessary in building our understanding of best practices in terrorism prevention efforts, but they also provide insight into potential benefits, hidden costs, and unintended consequences of targeted interventions, trainings, and community-focused efforts, with ramifications for national security policy.

Although it is essential to evaluate terrorism prevention programs, there are inherent difficulties in doing so. Such difficulties are not necessarily unique to terrorism prevention evaluation efforts, but they do present important challenges to those conducting them. Although evaluating the impact of some programs may be relatively straightforward in certain circumstances — for example, evaluating the effectiveness of smoking cessation efforts would entail measuring the extent to which individual smokers going through those programs quit smoking — evaluating the impact of terrorism prevention is more complex. This is in large part due to what is often deemed the “difficulty of measuring a negative.”⁹ In the case of terrorism prevention, how do we measure the terrorist attacks or terrorist recruitment that did *not* occur because of a program?

Additional measurement and data issues require evaluators to be careful in designing and implementing their data collection, analyzing data, and developing conclusions. These issues can include challenges associated with identifying measurable variables or proxy variables, controlling for social desirability bias, ensuring programmatic participation and evaluation response rates, and data availability.¹⁰ These difficulties and their impact on the lessons derived from NIJ-sponsored evaluations in this report are discussed below.

Although these issues make high-quality evaluation difficult, they are not necessarily insurmountable. Indeed, various mechanisms can be put in place to alleviate some of these challenges. NIJ-funded evaluation research has created many instructive and, importantly, accessible resources to ease the process of and comparability between evaluations of terrorism prevention programs. For example, NIJ-funded evaluation research on the role of peers in recognizing and reporting early signs of radicalization to terrorism led to the creation of an empirically tested, open-access set of measurable criteria that can be broadly used and easily adapted to standardize future program evaluations and research.¹¹ Additional NIJ-sponsored research has also created evaluative frameworks useful in replicating and evaluating similar programs¹² as well as research findings useful in creating, assessing, and altering programs focused on assisting individuals who are exiting extremism and reintegrating into their communities.¹³ This report synthesizes findings from NIJ-funded terrorism prevention evaluations and highlights additional lessons that may help ameliorate similar challenges in future evaluative efforts.

NIJ-Sponsored Evaluation Projects

In this section we briefly describe each project reviewed before examining three of those projects in greater depth.¹⁴ Exhibit 1 lists completed NIJ Domestic Radicalization to Terrorism-sponsored research focused on different forms of evaluation and assessment reviewed and coded by the authors of this report.

Exhibit 1: Overview of Reviewed NIJ-Sponsored Evaluations

Years	Grant No.	Project Title	Grantee	Program Type	Intended Audience	Overview
2012-2015	DJO-NIJ-12006GS-10F-0275P	Evaluation for State and Local Law Enforcement Anti-Terrorism Training (SLATT) Program	RAND Corporation	Training, Law Enforcement Terrorism Awareness	Law Enforcement	RAND Corporation evaluated the SLATT program, which provides counterterrorism training to state, local, and tribal law enforcement. ¹ This program is reviewed in depth below.
2012-2014	2013-ZA-BX-0003	Evaluation of a Multi-Faceted, U.S. Community-Based, Muslim-Led CVE Program	University of Massachusetts-Lowell	Peer Gatekeeping	Youth	The University of Massachusetts-Lowell team evaluated a countering violent extremism (CVE) program in Montgomery County, Maryland, led by the World Organization for Resource Development and Education. ² This evaluation informed the later project by ANSER, which is reviewed below.
2013-2017	2013-ZA-BX-0004	Evaluating the Federal CVE Initiative	Duke University	Policy Evaluation	Policy Community	Duke University examined the CVE Initiative (2011-2017). They surveyed U.S. attorneys and spoke to federal agents, law officials, and a Muslim-American focus group. ³
2016-2018	2015-ZA-BX-0002	Readiness Evaluation for Community Resilience	Research Triangle Institute	Workshops, Community and Law Enforcement Collaboration	Law Enforcement and Community	Research Triangle Institute reviewed documents and conducted interviews to evaluate the Community Resilience Exercise program, a half-day table-top exercise designed to help communities realize their violent extremism risks, identify available resources, and create a community action plan. ⁴

Exhibit 1: Overview of Reviewed NIJ-Sponsored Evaluations (continued)

Years	Grant No.	Project Title	Grantee	Program Type	Intended Audience	Overview
2016-2019	2015-ZA-BX-0003	Evaluating the Safe Spaces Community-Led CVE Program	University of Illinois at Chicago	Training, Community Terrorism Awareness, Prevention, and Intervention	Community (Mosques)	The University of Illinois at Chicago evaluated the Safe Spaces program for empowering Muslim communities to prevent violent actions. ⁵ This program is reviewed in depth below.
2017-2018	2016-ZA-BX-K001	Evaluation of the Peer to Peer (P2P): Challenging Extremism Initiative	President and Fellows of Harvard College/Harvard T.H. Chan	Online Terrorism Prevention Initiatives, Awareness and Education	Youth, Online	The Harvard team evaluated programs to educate youth on violent extremism and enhance tolerance created as part of the P2P Challenging Extremism Initiative. ⁶
2017-2018	2016-ZA-BX-K003	Using Gatekeeper Training as a CVE Tool: Replication and Evaluation of a Gatekeeper Program in Prince George's County	ANSER (Analytic Services, Inc.)	Training, Peer Gatekeeping, Terrorism Awareness	Youth	ANSER evaluated the Global Citizen's Forum, an after-school program on preventing violent extremism among youth through peer gatekeeping. ⁷ This program is reviewed in depth below.
2019-2021	2018-ZA-CX-0002	Operation 250: An Evaluation of a Primary Prevention Campaign Focused on Online Safety and Risk Assessment	University of Massachusetts-Lowell	Online Terrorism Prevention Initiatives, Awareness and Education	Youth, Online	The University of Massachusetts-Lowell partnered with Operation 250 (Op250) to conduct a formal and summative assessment using randomized controlled trials of Op250's online safety- and risk-focused counterterrorism programs for students. ⁸

- Lois M. Davis et al., "Assessment of the State and Local Anti-Terrorism Training (SLATT) Program," Final report to the National Institute of Justice, grant number DJO-NIJ-12006, 2016, NCJ 250418, <https://www.ojp.gov/pdffiles1/nij/grants/250418.pdf>.
- Michael J. Williams, John G. Horgan, and William P. Evans, "Evaluation of a Multi-Faceted, U.S. Community-Based, Muslim-Led CVE Program," Final report to the National Institute of Justice, grant number 2013-ZA-BX-0003, June 2016, NCJ 249936, <https://www.ncjrs.gov/pdffiles1/nij/grants/249936.pdf>.
- David Schanzer and Joe Eyerman, "Engaging With Communities To Prevent Violent Extremism: A Review of the Obama Administration's CVE Initiative," Final report to the National Institute of Justice, grant number 2013-ZA-BX-0004, January 2021, NCJ 256018, <https://www.ojp.gov/pdffiles1/nij/grants/256018.pdf>.
- RTI International, "Evaluability Assessment and Formative Review of the Community Resilience Exercises (CREX): Summary Overview," Final report to the National Institute of Justice, grant number 2015-ZA-BX-0002, January 2021, NCJ 256034, <https://www.ojp.gov/pdffiles1/nij/grants/256034.pdf>.
- Stevan Weine et al., "Evaluating the Safe Spaces Program: Using a Community-Based Public Health Approach To Prevent Violent Extremism," Final report to the National Institute of Justice, grant number 2015-ZA-BX-0003, January 2021, NCJ 256025, <https://www.ojp.gov/pdffiles1/nij/grants/256025.pdf>.
- Elena Savoia, "Evaluation of the Peer to Peer (P2P) Challenging Extremism Initiative," Final report to the National Institute of Justice, grant number 2016-ZA-BX-K001, August 2020, NCJ 255101, <https://www.ojp.gov/pdffiles1/nij/grants/255101.pdf>.
- Analytic Services Inc., "Using Gatekeeper Training as a CVE Tool: Replication and Evaluation of a Gatekeeper Program in Prince George's County," Final report to the National Institute of Justice, grant number 2016-ZA-BX-K003, January 2021, NCJ 256022, <https://www.ojp.gov/pdffiles1/nij/grants/256022.pdf>.
- National Institute of Justice funding award description, "Operation250: An Evaluation of a Primary Prevention Campaign focused on Online Safety and Risk Assessment," at the University of Massachusetts, award number 2018-ZA-CX-0002, <https://nij.ojp.gov/funding/awards/2018-za-cx-0002>.

Spotlight: An In-Depth Overview of Select NIJ-Sponsored Evaluations

NIJ-sponsored evaluations took place at different times, in different communities and locations, with different programmatic activities and goals, and using different methods of analysis. This makes it difficult to draw generalizable lessons from across the portfolio. In an effort to imbue a level of comparability in synthesizing findings, we selected three

of the evaluations with sufficient information to allow for a detailed review of findings across similar programs implemented for different audiences: (1) law enforcement, (2) communities, and (3) youth. These three grants are expanded upon in further detail to highlight some of the prevalent characteristics the programs evaluated; issues associated with the implementation of the evaluation, including details about evaluation modification and limitations; and evaluation findings and subsequent implications.

As noted previously, the evaluations themselves should not be viewed as directly comparable in a scientific sense. However, examining evaluations focused on similar types of programming, but implemented for different audiences, is advantageous for several reasons. Importantly, it provides an understanding of the comparability across evaluations with different audiences and the types of similarities and differences that may be observed. In the absence of generalizable findings from the evaluations themselves, this loose comparability provides some insight into whether the lessons derived from each evaluation may be useful or even relevant to similar programming for different communities. In addition, it provides insight into the efficacy of and challenges faced in conducting similar programs (in this case training, education, and awareness-related efforts) and subsequent evaluations of them.

Finally, since the implementation of these evaluations (all occurring between 2012 and 2019), research on evaluation best practices, violent extremism, and terrorism has expanded. Although these evaluations may not reflect the current state of thinking on terrorism prevention, assessment, and evaluation, they do provide instructive lessons on issues the field continues to grapple with today.

Evaluating Law Enforcement-Focused Programs: Assessment of the State and Local Law Enforcement Anti-Terrorism Training (SLATT) Program (2012-2015)

In their project “Evaluation for State and Local Law Enforcement Anti-Terrorism Training (SLATT) Program,” RAND Corporation (RAND) conducted an assessment of the SLATT program, a workshop-based training program created by the Bureau of Justice Assistance. The program provides state, local, and tribal law enforcement with instruction on key topics of interest, including training on understanding, detecting, and responding to acts of terrorism and violent extremism from both international and domestic actors and groups.¹⁵ SLATT also provides law enforcement with access to an online database with relevant materials.

RAND’s assessment of SLATT focused on two types of workshop activities: (1) investigative/intelligence workshops designed for law enforcement personnel interested in learning more about mechanisms for identifying and addressing terrorism threats and (2) train-the-trainer workshops designed for law enforcement personnel interested in crafting and implementing their own training. The RAND evaluators surveyed law enforcement personnel who participated in either type of trainings to answer questions about how each type of training was carried out; who participated; the costs, benefits, and challenges associated with participation; the impact of the trainings; and how the trainings could be improved.¹⁶

The project team utilized a mixed-methods approach involving a literature review, program observation, interviews with participants and program staff, a follow-up online participant survey (including discrete choice experiments to measure the perceived value

of the program among participants), and the examination of website user and engagement metrics.¹⁷ Survey data were gathered from participants in two investigative/intelligence workshops and three train-the-trainer workshops implemented between late 2014 and early 2015.¹⁸ Collected data were also disaggregated to explore differences between experiences and perceptions in the two types of workshops and variations among the participants in each workshop.

The RAND team was clear about the limitations to their assessment and its implications for the generalizability of their findings. First, because SLATT workshops were implemented upon request from local law enforcement, the team was unable to achieve the desired variation in training focus, type, and location, making it impossible to determine whether the aspects of training they did observe were representative of SLATT training overall. Second, it was not possible to establish a control or comparison group, which limits the overall understanding of nonparticipant law enforcement training needs, potential benefits, and perceptions. Finally, RAND's assessment regarding if and how the training altered law enforcement approaches to terrorism prevention was limited to self-reported data that they were unable to independently verify.¹⁹

Despite these limitations, findings from the assessment indicate some important lessons for similar efforts. The evaluation noted that the training seemed to attract the desired audiences (according to the research team, possibly due to the self-selection of participants into a program they were already interested in attending).²⁰ Participants also expressed interest in the topics and reported that the workshops' information on international and domestic terrorist threats was useful. Investigative/intelligence workshop participants expressed higher ratings for usefulness than train-the-trainer workshop participants,²¹ which would suggest the trainings themselves are seen as relevant among law enforcement audiences working on counterterrorism issues.

Overall, most participants (65%) agreed that information from the workshops would affect how they investigated and approached domestic and international terrorist threats.²² The evaluation team was not able to assess the extent to which these claims manifested in actual behavioral changes, however. It should also be noted that no baseline data on participant knowledge and motivations were collected prior to implementing the trainings, as doing so was outside the scope of the evaluation effort. Still, respondents from both workshops indicated continued motivation to attend SLATT training based on RAND's cost-benefit analysis, with participants from the investigative/intelligence workshops indicating a greater desire for the inclusion of concrete terrorist threat examples specific to their own locality and jurisdiction.²⁴

Although there are limitations to generalizing too broadly from the findings, given the nature of the evaluation and methods used, RAND's assessment of SLATT provides useful findings and lessons for further programmatic understanding and assessment. Notably, distinctions between project participants (law enforcement officers and command staff/supervisors vs. prosecutors, regional planners, etc.) in both workshops (investigative intelligence vs. train-the-trainer workshops) may indicate the need to further refine the training provided based on the needs, level of expertise, and expectations of the participants in each workshop, which varied by group.²⁵ "Train-the-trainer workshop participants differed in some important ways from trainees who attended the SLATT intelligence/investigative workshops," as the former tended to be: (1) farther along in

their careers, (2) more likely to travel a longer distance to attend the workshops, (3) more proactive in professional networking at the events, and (4) planned to develop their own training classes based on the workshop material.²⁶ Although participants noted overall that the information provided by the training was useful, without further assessment of how participants used the concepts and information from the training, and how they compare to participant knowledge and expectations prior to the trainings, the evaluation was unable to determine the training's impact and efficacy.

At the time of writing, the SLATT program is still active. In-person trainings were briefly suspended due to the COVID-19 pandemic, but virtual options are now available for some training types.²⁷ In addition, the SLATT website incorporates a number of e-learning courses and resources, ranging from webinars to documents and podcasts, that are accessible to law enforcement officers. These resources cover a variety of topics that may address different interests and expectations of participants noted during the evaluation in a customizable and nuanced manner.²⁸ Additional evaluative efforts, including baseline assessments, would be useful in determining the extent to which these resources address the original evaluation findings, and in assessing whether any changes to the curriculum and training adopted since the evaluation have impacted law enforcement practices.

Community-Focused Programming: Evaluating the Safe Spaces Community-Led Countering Violent Extremism (CVE) Program (2016-2019)

In “Evaluating the Safe Spaces Community-Led CVE Program,” the University of Illinois at Chicago partnered with the Muslim Public Affairs Council (MPAC) to modify and subsequently evaluate the implementation and impact of Safe Spaces, a community-led public health and violence prevention model for preventing violent extremism.²⁹ The Safe Spaces program was originally developed by MPAC to “empower Muslim communities to protect themselves against misguided ideas and violent actions through community investment in social and religious programs and community adoption of multidisciplinary ‘crisis inquiry teams’ that intervene when an individual is at risk of engaging in violent behavior.”³⁰ The NIJ-sponsored evaluation of the Safe Spaces program focused on: (1) training external trainers to implement a curriculum that provides instructive education on the Safe Spaces model, including its prevention and intervention components; (2) delivering the curriculum to participants from local mosques in nine different implementation sites; and (3) encouraging trainee development of and participation in follow-on prevention and intervention activities.

Adopting a public health framework, the program incorporated a bottom-up approach, consisting of both prevention- and intervention-focused layers. During the program modification phase prior to implementation, feedback from community focus groups led evaluators to eliminate another layer — ejection. According to the project team, “ejection” refers to “removing a person from their organization” in the PIE model of Prevention, Intervention, and Ejection. Explicit focuses on violent extremism and national security and references to individual terrorism risk factors were also removed, along with ejection, during the redesign effort in favor of community public health measures and violence prevention more broadly (which were identified as broader needs of the community).

The evaluation of the training and follow-on prevention- and intervention-focused activities initially sought to use survey and quantitative analysis to understand (1) how

community-based organizations develop effective CVE programs; (2) the evaluation tools and procedures that community organizations, law enforcement, and government agencies can use to build effective CVE programs; and (3) whether prevention and intervention activities conducted through Safe Spaces diminish risk factors or enhance protective factors.³¹

Ultimately, difficulties associated with program implementation necessitated scaling back the evaluative focus, sites, and measures, with a notable impact on the project team's ability to answer the original questions posed. Lack of participant engagement also presented problems for the evaluators; many participants attended only some sessions, while others were not sufficiently aware of the associated prevention and intervention activities and expected commitments following the training. Out of the nine sites, only one continued with both prevention and intervention efforts following the training. Therefore, the team focused the evaluation on process measures — implementation issues, participant reflections on the program, and recommendations for future efforts — through follow-up qualitative interviews.³²

The findings indicated that, overall, the Safe Spaces program was not successful in reaching its intended outcomes and impact but highlighted important areas of opportunity and consideration. First among those was the importance of scoping and understanding individual community needs and engaging with community leaders *prior* to designing and implementing the program. In this case, unlike the SLATT law enforcement trainings described above, community needs and concerns were less related to terrorism.³³ According to the evaluation team, prior engagement helps clarify the nature, scope, and purpose of programming; it also can help ensure buy-in from local community leaders, something that was not present across all nine community sites.

Although the evaluation team attempted to scope out community needs during the modification phase, ultimately they found these efforts to be insufficient, impacting implementation of the program overall. Not only did the communities participating in the training express a lack of concern about violent extremism and terrorism within their own communities, they also found the program's general focus on violence to be unrepresentative of their needs and stigmatizing, given concerns about profiling and the national security focus of terrorism prevention efforts.³⁴ Although the training and program were subsequently altered to focus on public health, the evaluators found that the redesigned program was duplicative to some extent, with participants reporting that similar activities were already occurring within their communities.

Second among these findings was the importance of clarification, cohesive curricula, and trainer expertise. Participants noted confusion regarding the purpose, aims, and scope of the program. As the evaluators acknowledged, some of this confusion may have resulted from participants' lack of knowledge about the redesigned program as well as what participants noted as a lack of knowledge on the part of the trained trainer. The evaluators noted that attempting to meld public health and violence prevention or threat assessment efforts led to confusion.

Third among these findings is the need for dedicated resources and support post-training. Although the training provided information relevant to the development of further prevention and intervention initiatives, without the funding, expertise, and buy-in to conduct the prevention and intervention initiatives following the initial training, the

program faced difficulties in implementing these activities. Indeed, only one group of trainees was able to conduct follow-on prevention and intervention activities, and only with dedicated support and funding from the evaluation team.

In summary, although challenges associated with program implementation necessitated alterations in the original evaluation design — thereby limiting the rigor and scope of the evaluation effort — the researchers were able to highlight important aspects to consider in implementing public health-based terrorism prevention efforts at the community level. Notably, these included the importance of implementing such programs in a variety of communities that are concerned about violence, rather than focusing programmatic efforts on specific ethnic or religious communities, which can lead to stigmatization.³⁵ These findings and recommendations suggest an opportunity for further efforts to determine whether and how terrorism prevention might be successfully incorporated within broader violence prevention and community resilience-focused efforts.

Evaluating Youth-Focused Programs: The Global Citizen’s Forum (2017-2018)

In their project “Using Gatekeeper Training as a CVE Tool: Replication and Evaluation of a Gatekeeper Program in Prince George’s County,” the research team from Analytic Services, Inc. (ANSER) conducted an evaluation of the Global Citizen’s Forum (GCF). An after-school program focused on preventing violent extremism among youth through peer gatekeeping and awareness efforts, the GCF was developed by the World Organization for Resource Development and Education.³⁶ The program sought to empower “youth to identify and report possible signs of radicalization to terrorism” through a curriculum designed to teach youth (ages 14-17) to recognize signs of radicalization and identify avenues for further support if they witness those signs in their peers.³⁷

The evaluation of the GCF focused on: (1) replicating and piloting the GCF peer gatekeeping model and curriculum in a new location — Prince George’s County, Maryland; (2) training trainers to deliver the GCF to youth in after-school programs; and (3) assessing and evaluating the extent to which the GCF model could be replicated, its effectiveness in addressing violent extremism among youth, and its cost-effectiveness. ANSER’s approach to evaluating the program consisted primarily of gauging the short-term impact on attitudes and knowledge among youth participants.³⁸

The evaluation team originally planned to implement quasi-experimental methods, qualitative analyses, and cost analyses to assess the program’s impact in increasing understanding of violent extremism among youth, the efficacy of gatekeeper programs in addressing violent extremism, and the replicability of the model. However, due to issues associated with insufficient sample size, program attrition, unexpected developments and barriers to working within school systems, and low survey response rates and quality, the researchers were not able to obtain a sample size sufficient to fully evaluate the program using more rigorous statistical measures. These factors necessitated scaling back the evaluative focus, measures, and design to focus on “exploratory research” through limited survey responses and focus group discussions.³⁹

Despite these challenges and limitations, findings from the evaluation point to important factors related to program implementation, impact, and potential for further replication.⁴⁰ As with the Safe Spaces program described above, the research team found that sensitivities

to the term “violent extremism” and lack of a perceived need for violent extremism prevention programming raised concerns among both program participants and school administrators. The researchers noted that explicit focus on violent extremism may serve as an impediment to program implementation and suggested that focusing on violence prevention (instead of violent extremism) may prove useful in alleviating implementation issues. Not only was violent extremism viewed as sensitive, necessitating further approval processes, but it also was not necessarily seen as an important topic among school officials because they did not view terrorism as a threat within their communities. Among participants in the GCF program, 50% (six of 12) noted in post-implementation survey responses that they disagreed or strongly disagreed that violent extremism was a problem within their communities.

Despite this, the team noted positive findings related to the program. Most of the respondents agreed or strongly agreed that they found the training interesting or informative and would recommend it to their peers. Beyond terrorism prevention outcomes, respondents reported that the program improved general social and public speaking skills, while some noted that they planned to or had already used the concepts from the training to address circumstances unrelated to terrorism or violent extremism. Despite an initial scaled-back comparative analysis indicating a modest shift in participant perceptions of knowledge regarding violent extremism and attitudes toward gatekeeping, the evaluation ultimately fell short in effectively testing participants’ capacity to fully understand or address issues related to violent extremism among their peers, due to the limited sample size.⁴¹

ANSER’s evaluation of youth-focused GCF peer gatekeeping programs was not able to establish the GCF’s suitability in achieving terrorism prevention outcomes. Although this was due in part to challenges associated with program implementation that impacted evaluation design, rigor, and scope, further evaluative efforts — and possibly program modifications — are recommended.

Synthesized Lessons: Program- and Evaluation-Specific Lessons From NIJ Evaluation Efforts

Although variations in the evaluation methods, types, and programmatic focuses exist within the NIJ-sponsored evaluative research, a few key lessons and insights related to both programmatic outcomes and evaluation practices emerged. Importantly, despite the varied focus of the evaluations analyzed, some of them shared similar lessons regarding both programmatic activities and evaluation processes, suggesting these lessons may be more universal to terrorism programs generally, or at least to those that are similar in nature and audience.

Program-Specific Lessons

Gaining Community Buy-In and Assuring Program Relevance for Participants

Overall, lessons from NIJ-sponsored terrorism prevention evaluations suggest more specific attention is needed to ensure that terrorism prevention efforts are crafted in consultation and cooperation with the audiences for which they are constructed, so that they are (1) accepted by and seen as an added value to those audiences and (2) appropriately tailored to individual audiences' concerns, desires, and needs.

Given the national security focus of such programs, it is additionally important that audience consultation take place to address concerns regarding the stigmatization these programs can entail — especially within community-focused programs, where community characteristics may lead participants to feel targeted or singled out as national security threats.⁴² In addition, it is important to gauge expectations and concerns about terrorist threats among participants before implementing and designing terrorism prevention initiatives. ANSER's findings from the youth-focused GCF program indicated that not only was violent extremism not widely viewed as a concern among youth participants in

the program, but the focus on violent extremism also led to skepticism and pushback from administrators in the schools where the program was meant to be implemented.⁴³ Meanwhile, RAND's assessment of SLATT workshops indicated that even within a program intended for counterterrorism practitioners and law enforcement, audience priorities around training of terrorism-related issues may differ, with potential implications for program impact and efficacy. This lesson, shared across evaluations of three diverse programs, highlights the importance of initial scoping efforts to determine program fit and desirability before implementation; such efforts have the potential to increase the efficacy, participant buy-in, and sustainability of terrorism prevention programs and avoid negative outcomes.

Programmatic Impact Beyond Terrorism Prevention

Although it may be difficult to assess the actual impact of programs on terrorism prevention goals, findings across the evaluations pointed to the relevance of non-terrorism prevention outcomes and benefits from evaluated terrorism prevention programs. Findings from the GCF, for example, indicated that even though youth participants did not believe that violent extremism was a matter of concern within their community, they used, or planned to use, the information from the program to address issues unrelated to violent extremism and experienced benefits from the program unrelated to terrorism prevention outcomes.⁴⁴ Participants in SLATT training also noted benefits beyond terrorism prevention instruction, including expanding their professional networks.⁴⁵ Although not generalizable to terrorism prevention broadly, this may suggest that terrorism prevention programs have unintended benefits beyond simply preventing terrorism, with potential implications for broader social goals. Future evaluations should assess impact both directly related to and beyond simple terrorism prevention outcomes to assess any auxiliary benefits and determine whether programs providing those benefits without a focus on terrorism prevention might be more suitable to audience needs.

Incorporating Terrorism Prevention in Broader Violence Prevention and Public Health Programming

Finally, findings from the programs shed additional light on a long-standing topic of discussion in terrorism prevention spaces: How explicit should the focus on terrorism prevention be if it may have the effect of alienating or stigmatizing the communities in which a program intends to lend support? Findings from across grants were mixed in this regard, noting the benefits of including terrorism prevention within wider violence prevention and public health initiatives while also noting the difficulties experienced in attempting to do so.⁴⁶

Findings from projects with youth and community audiences indicated that “violent extremism” was not necessarily a popular term, nor was it necessarily identified as an area of concern among the individuals who participated in programmatic activities. In the case of evaluations for Safe Spaces and the GCF, for example, violence prevention, public health, and community resilience seemed to be more commonly used and widely acceptable terms for program types among those participating and were suggested by the project teams as potentially more viable names for terrorism prevention programs.

However, when the main intent of a program is terrorism prevention specifically, this broader focus can dilute efforts to measure the outcomes and impact of the project and confuse participants, especially when elements of the project are still linked to terrorism

prevention.⁴⁷ Recommendations for how to better incorporate terrorism prevention within broader violence prevention and community health models were specifically noted, including the use of public health terminology in redesigning programs. Further evaluation of programs that fully incorporate those recommendations should be considered before drawing any conclusions about their real or potential efficacy.

Evaluation-Specific Lessons

As noted earlier, many of the findings above — and the evaluations themselves — were limited, by either design or necessity. Variations to initial evaluation designs were common, given challenges encountered upon implementation. Rather than disregard these shortcomings as failures, there are lessons to be learned from them along with potential findings that may help limit similar issues facing evaluators in the future.

Assessing Data Needs

The NIJ-sponsored evaluations reviewed in this report highlight the importance of conducting formative evaluations or evaluability assessments prior to program implementation and evaluation. Evaluability assessments can help identify data gaps, appropriate methods, and possible challenges before implementing and evaluating a program. Indeed, these are essential measures, given that lack of such data can impact results, lead to the collection of data ill-suited to evaluation goals, and impact the quality of findings.

Addressing Participant Attrition

Program attrition rates impacted evaluation plans and the ability to collect sufficient data for more robust statistical testing and verification.⁴⁸ Evaluations and programs should work together to determine how to retain program participants. Participant attrition rates significantly impacted efforts to evaluate the effect of these programs scientifically, resulting in limitations in the findings and their generalizability to individuals beyond those who initially responded. Evaluators should consider why individuals may not want to participate, incentives for participation, and adequate, long-term resources needed to ensure continued engagement in terrorism prevention initiatives and reduce participant attrition.

Measurement Strategy and Social Desirability Bias

The lack of controls for social desirability bias and measures to confirm the veracity of respondents' answers to evaluation prompts and questions causes doubts about the extent to which the findings related to individual actions, beliefs, and behaviors conform with reality.

Causal Identification Strategy

The absence of comparison measures and assessments in many of the evaluations makes it difficult to fully assess the extent to which stated outcomes can be attributed to programmatic activities. In an ideal world, each evaluation would include some type of control or comparison group. Where this is not feasible, baseline information to assess attitudes, behaviors, and knowledge prior to the implementation of programming would give us more confidence in determining whether the post-implementation measures are a result of the program. Moreover, the lack of follow-up assessments in many cases makes

it difficult to gauge long-term programmatic impacts on participants well after they have participated in programming. This makes it difficult to track whether programmatic activities might have actually led to behavioral, knowledge, or attitudinal changes based on prevention activities.

Increasing Evaluation Generalizability

Even the best evaluation has limitations in generalizing the findings beyond the individual projects. Many of the evaluations examined here specifically noted and discussed these limitations, which arose from issues associated with the inability to establish a treatment and control or comparison group, unexpected program attrition, or the inability to randomize the samples from which the findings were drawn.⁴⁹ Some evaluations either attempted to include significant safeguards — as was the case with the GCF and SLATT, both of which were ultimately unable to establish comparison and control groups due to program attrition or programmatic constraints — or suggested means by which future evaluations and research could improve issues associated with the generalizability of findings, such as through program and evaluation replication in other communities or environments.⁵⁰ Such efforts to draw attention to these limitations should be commended. Other evaluation efforts sponsored by NIJ notably sought to employ randomized controlled trials or strong quasi-experimental designs to increase the validity and generalizability of study findings.⁵¹ Given these difficulties and limitations, researchers should take caution in overgeneralizing the findings and lessons detailed in this report.

Gaps and the Path Forward

Along with important lessons, significant gaps emerged in this set of evaluation projects that are worthy of further funding and research. Although many of these gaps arise from difficulties in evaluating any type of program, others are more specific to the terrorism prevention field.

Evaluating Long-Term Impact

There is a need for further follow-up assessments to evaluate program effectiveness, impact, and knowledge retention well after participation in the initial program. Follow-up assessments that track reported or real changes from programmatic interventions are necessary, especially to examine impact and utility over time. Along with pre- and post-intervention assessments, follow-up assessments should be incorporated into future evaluation efforts to gauge knowledge retention and potential behavioral outcomes associated with programming.

Focusing on Behavioral Changes

Although NIJ-funded evaluations provided important insights into participant perceptions and attitudes toward the programs themselves, information about their actual impact on terrorism prevention-relevant behaviors among program participants was not as widely measured. To be sure, this is a difficult task in any evaluative effort. Even among evaluations that originally intended to measure behavioral outcomes, lack of data or appropriate pre-, post-, and follow-up measures, metrics, and methods prevented them from doing so. Greater attention to examining how programs affect behavioral outcomes or changes, while difficult, is necessary for understanding the intended impact, especially given the national security interests inherent in terrorism prevention policies and activities. Although no method is perfect, including follow-up assessments with specific questions gauging behavioral change and examples of such change, or incorporating behavior-based methods,⁵² may be beneficial in efforts to better approximate behavioral change.

Broadening the Ideological and Nonideological Focus of Program Evaluations

It is important to note that the NIJ-sponsored evaluations reviewed in this report did not focus on assessing the programs' impact on specific strands of terrorism (defined ideologically or otherwise), nor did they necessarily assess the extent to which programs might be equipped to prevent different types of terrorism. This is significant because the efficacy of terrorism prevention efforts in addressing different forms of terrorism remains unclear without further study. The evaluations assessed here were not able to determine the impact on terrorism prevention outcomes, nor were they able to determine the programs' effectiveness in preventing different types of terrorism threats. It is unclear how these programs and the findings from their evaluations are generalizable across different types of ideologically motivated radicalization to terrorism. Future evaluations should assess the extent to which assumptions regarding the suitability of terrorism prevention activities are sufficient in addressing varied ideological terrorism motivations.

Conclusion and Policy Implications

NIJ-sponsored evaluations of terrorism prevention programs have yielded important lessons for both evaluators and program designers. Although definitive conclusions cannot be drawn regarding the impact of these programs due to methodological constraints, overall the lessons and gaps identified in the research suggest certain implications for terrorism prevention evaluations and policy moving forward.

First, findings from NIJ-sponsored evaluations point to the importance of considering the appropriate approach to terrorism prevention. Although none of the evaluations were able to determine the impact of programming in preventing terrorism, some suggested the possible utility of including terrorism prevention within a broader portfolio of violence prevention and public health programs. The findings from these evaluations are limited; however, they point to difficulties in conceptualizing violent extremism and a lack of participant concern about violent extremism as impacting programmatic activities.⁵³ Although this may or may not apply to law enforcement programs or other programs focused specifically on counterterrorism practitioners, the benefits and potential issues should be explored further.

Second, greater focus should be placed on adopting clear definitions, goals, and strategies for terrorism prevention programming. In the absence of clear guidance on expected measures of impact and programmatic efforts, program developers and evaluators may not have the necessary information to construct and implement programmatic initiatives and evaluations to achieve desired policy results. Moreover, in the absence of clear definitions of terminology (e.g., violent extremism), it can be difficult to ensure that implemented programs are able to address the needs of terrorism prevention policy and communicate those needs clearly with program participants.

Finally, additional resources should be considered to ensure that the necessary expertise and financial resources are available for law enforcement and community-led terrorism prevention programs. Moreover, additional resources should be directed to evaluators pursuing long-term, methodologically sound evaluations focused on the impact of terrorism

prevention programs in achieving their goals. Although such efforts are notoriously difficult, further exploration of innovative methods and efforts can contribute to the valuable insights and lessons detailed here from NIJ-supported domestic terrorism prevention evaluation efforts.

Notes

1. For more information on this effort, see Aisha Javed Qureshi, “Understanding Domestic Radicalization and Terrorism: A National Issue Within a Global Context,” *NIJ Journal* 282, August 2020, <https://nij.ojp.gov/library/publications/understanding-domestic-radicalization-and-terrorism-national-issue-within>.
2. Allison G. Smith, “How Radicalization to Terrorism Occurs in the United States: What Research Sponsored by the National Institute of Justice Tells Us,” Washington, DC: U.S. Department of Justice, National Institute of Justice, 2018, <https://www.ojp.gov/pdffiles1/nij/250171.pdf>; and Allison G. Smith, “Risk Factors and Indicators Associated With Radicalization to Terrorism in the United States: What Research Sponsored by the National Institute of Justice Tells Us,” Washington, DC: U.S. Department of Justice, National Institute of Justice, 2018, <https://www.ojp.gov/pdffiles1/nij/251789.pdf>.
3. This report is the first of four reports based on NIJ research funded in the Domestic Radicalization and Terrorism portfolio between fiscal years 2015 and 2019 along with selected research projects from NIJ’s Hate and Bias Crimes and Mass Shootings portfolio. This report does not include forthcoming findings from ongoing NIJ-sponsored terrorism prevention assessment and evaluation efforts.
4. Evaluations from these years were spotlighted based on their completion, availability of data, and similar programmatic foci. It is worth noting that since the implementation of these evaluations, the field and nature of terrorism and terrorism prevention efforts have grown notably. Lessons drawn from evaluations included in this report, therefore, are not representative of the current state of knowledge or practice. Nevertheless, they provide important insights and snapshots of the field of evaluation and practice during that time that may remain relevant to present-day efforts.

5. Georgia Holmer, Peter Bauman, and Kateira Aryaeinejad, “Measuring Up: Evaluating the Impact of P/CVE Programs,” Washington, DC: U.S. Institute of Peace, 2018.
6. Evaluability assessments and formative evaluations both serve as a type of “pre-evaluation” evaluation. They focus on evaluating the extent to which a program is ready or developed sufficiently for evaluation and on providing information helpful in improving or modifying program delivery. This includes whether a program is feasible in its goals and outputs, reaches the right audiences, and is developed to the extent to which measurement is possible. For more information, see Caitlin Mastroe and Susan Szmania, “Surveying CVE Metrics in Prevention, Disengagement and Deradicalization Programs,” College Park, MD: National Consortium for the Study of Terrorism and Responses to Terrorism, 2016; and Peter Romaniuk and Naureen Chowdhury Fink, “From Input to Impact: Evaluating Terrorism Prevention Programs,” Center on Global Counterterrorism Cooperation, 2012.
7. Process evaluations focus on program implementation, including the extent to which activities adhere to original program designs, challenges encountered, and best practices for carrying out activities, rather than the impact of those activities in achieving stated outcomes and goals. Process evaluations may also examine the extent of participant satisfaction with programmatic activities. For more information, see Romaniuk and Chowdhury Fink, “From Input to Impact.”
8. Outcome and impact evaluations seek to measure the effect of a program, or how programmatic activities achieved (or did not achieve) overarching goals and outcomes such as changes in participant behaviors and attitudes. For example, an evaluation may look at how participants’ support for acts of terrorism before a program compares to their support after program participation or compared to a control group of nonparticipants. For more information, see Holmer, Bowman, and Aryaeinejad, “Measuring Up.”
9. Romaniuk and Chowdhury Fink, “From Input to Impact.”
10. Mastroe and Szmania, “Surveying CVE Metrics in Prevention, Disengagement and Deradicalization Programs.”
11. Michael J. Williams, John G. Horgan, and William P. Evans, “Evaluation of a Multi-Faceted, U.S. Community-Based, Muslim-Led CVE Program,” Final report to the National Institute of Justice, grant number 2013-ZA-BX-0003, June 2016, NCJ 249936, <https://www.ncjrs.gov/pdffiles1/nij/grants/249936.pdf>.
12. Analytic Services Inc., “Using Gatekeeper Training as a CVE Tool: Replication and Evaluation of a Gatekeeper Program in Prince George’s County,” Final report to the National Institute of Justice, grant number 2016-ZA-BX-K003, January 2021, NCJ 256022, <https://www.ojp.gov/pdffiles1/nij/grants/256022.pdf>.
13. Michael Jensen and Pete Simi, “Empirical Assessment of Domestic Disengagement and Deradicalization (EAD),” Final report to the National Institute of Justice, grant number 2014-ZA-BX-0003, January 2021, NCJ 256039, <https://www.ojp.gov/pdffiles1/nij/grants/256039.pdf>.

14. Additional NIJ grants focused on evaluations and assessments of terrorism prevention activities are ongoing and thus are not listed or included in this report. They include “Understanding the Potential for Multidisciplinary Threat Assessment Teams To Prevent Terrorism: Conducting a Formative Evaluation of the MassBay Threat Assessment Team,” grant number 2020-ZA-CX-0002, <https://nij.ojp.gov/funding/awards/2020-za-cx-0002>; “Evaluability Assessment and Development of Psychological and Behavioral Health Approaches To Prevent Terrorism and Facilitate Reintegration of Violent Extremists,” grant number 15PNIJ-21-GG-02727-DOMR, <https://nij.ojp.gov/funding/awards/15pnij-21-gg-02727-domr>; and “Towards Understanding Deradicalization in the U.S.: A Formative Evaluation and Evaluability Assessment of Parents for Peace,” grant number 15PNIJ-21-GG-02724-DOMR, <https://nij.ojp.gov/funding/awards/15pnij-21-gg-02724-domr>.
15. Lois M. Davis et al., “Assessment of the State and Local Anti-Terrorism Training (SLATT) Program,” Final report to the National Institute of Justice, grant number DJO-NIJ-12006, 2016, NCJ 250418, <https://www.ojp.gov/pdffiles1/nij/grants/250418.pdf>.
16. Davis et al., “Assessment of the State and Local Anti-Terrorism Training (SLATT) Program.”
17. Davis et al., “Assessment of the State and Local Anti-Terrorism Training (SLATT) Program.”
18. According to the research team, 89 out of 169 completed surveys were received from Wave I, while 27 out of 57 completed surveys were received from Wave II.
19. Davis et al., “Assessment of the State and Local Anti-Terrorism Training (SLATT) Program.”
20. Most participants reported counterterrorism as a primary responsibility within their positions.
21. Davis et al., “Assessment of the State and Local Anti-Terrorism Training (SLATT) Program.”
22. Davis et al., “Assessment of the State and Local Anti-Terrorism Training (SLATT) Program.”
23. Davis et al., “Assessment of the State and Local Anti-Terrorism Training (SLATT) Program.”
24. However, the caveat is that investigative/intelligence workshop participants expressed higher interest in receiving locally based examples than did participants in the train-the-trainer workshops.
25. Davis et al., “Assessment of the State and Local Anti-Terrorism Training (SLATT) Program.”
26. Davis et al., “Assessment of the State and Local Anti-Terrorism Training (SLATT) Program.”
27. According to the SLATT training website page, “Due to the current health crisis, our on-site trainings are postponed.” SLATT, “Training,” <https://www.slatt.org/Training#protect>, accessed May 25, 2022. As of November 14, 2022, it appears training activities may be resuming, albeit in a virtual format.
28. The authors of this report were unable to access the upload dates for these resources and were unable to access the content because of the gated nature of the website. For more information on the resources offered, see SLATT, “Resources,” <https://www.slatt.org/Resources>, accessed May 25, 2022.

29. Stevan Weine et al., “Evaluating the Safe Spaces Program: Using a Community-Based Public Health Approach To Prevent Violent Extremism,” Final report to the National Institute of Justice, grant number 2015-ZA-BX-0003, January 2021, NCJ 256025, <https://www.ojp.gov/pdffiles1/nij/grants/256025.pdf>.
30. Stevan Weine et al., “How Can a Public Health Framework Be Applied to Preventing Violent Extremism?” Final report to the National Institute of Justice, grant number 2015-ZA-BX-0003, January 2021, NCJ 256026, <https://www.ojp.gov/pdffiles1/nij/grants/256026.pdf>.
31. Weine et al., “Evaluating the Safe Spaces Program.”
32. Weine et al., “Evaluating the Safe Spaces Program.”
33. Community-identified concerns consisted largely of issues associated with social and familial dynamics, including mental health, substance abuse, youth decisions to leave their religion, etc.
34. Weine et al., “Evaluating the Safe Spaces Program”; and Analytic Services Inc., “Using Gatekeeper Training as a CVE Tool.” For more information, see also David Schanzer and Joe Eyerma, “Engaging With Communities To Prevent Violent Extremism: A Review of the Obama Administration’s CVE Initiative,” Final report to the National Institute of Justice, grant number 2013-ZA-BX-0004, January 2021, NCJ 256018, <https://www.ojp.gov/pdffiles1/nij/grants/256018.pdf>.
35. Weine et al., “Evaluating the Safe Spaces Program.”
36. This program was developed in response to findings from previous NIJ-sponsored evaluation efforts. See Williams, Horgan, and Evans, “Evaluation of a Multi-Faceted, U.S. Community-Based, Muslim-Led CVE Program.”
37. Analytic Services Inc., “Using Gatekeeper Training as a CVE Tool.”
38. Analytic Services Inc., “Using Gatekeeper Training as a CVE Tool.”
39. Analytic Services Inc., “Using Gatekeeper Training as a CVE Tool.”
40. Notable among these evaluation outputs was the creation of a replication and evaluation handbook for future efforts incorporating lessons from this evaluation. See Analytic Services Inc., “The Global Citizen’s Forum (GCF) Replication Framework,” Final report to the National Institute of Justice, grant number 2016-ZA-BX-K003, January 2021, NCJ 256023, <https://www.ojp.gov/pdffiles1/nij/grants/256023.pdf>.
41. Analytic Services Inc., “Using Gatekeeper Training as a CVE Tool.”
42. See, for example, the previous discussion of the Safe Spaces program.
43. Analytic Services Inc., “Using Gatekeeper Training as a CVE Tool.”
44. Analytic Services Inc., “Using Gatekeeper Training as a CVE Tool.”

45. Davis et al., “Assessment of the State and Local Anti-Terrorism Training (SLATT) Program.”
46. For example, Analytic Services Inc., “Using Gatekeeper Training as a CVE Tool”; and Weine et al., “Evaluating the Safe Spaces Program.”
47. Weine et al., “How Can a Public Health Framework Be Applied to Preventing Violent Extremism?”; Weine et al., “What Should Program Designers Consider to Successfully Develop and Implement a Public Health Approach to Preventing Violent Extremism?,” Final report to the National Institute of Justice, grant number 2015-ZA-BX-0003, January 2021, NCJ 256027, <https://www.ojp.gov/pdffiles1/nij/grants/256027.pdf>; and Weine et al., “What Does a Community-Based Organization Need To Successfully Implement a Public Health Approach to Preventing Violent Extremism?,” Final report to the National Institute of Justice, grant number 2015-ZA-BX-0003, January 2021, NCJ 256028, <https://www.ojp.gov/pdffiles1/nij/grants/256028.pdf>.
48. Analytic Services Inc., “Using Gatekeeper Training as a CVE Tool”; and Weine et al., “Evaluating the Safe Spaces Program.”
49. Analytic Services Inc., “Using Gatekeeper Training as a CVE Tool”; Davis et al., “Assessment of the State and Local Anti-Terrorism Training (SLATT) Program”; and Weine et al., “Evaluating the Safe Spaces Program.”
50. Analytic Services Inc., “Using Gatekeeper Training as a CVE Tool.”
51. They include NIJ-sponsored evaluations of the P2P and Op250 programs, both of which will be examined in future reports.
52. Where appropriate, lab-in-field experiments represent one such potential method. Although they have limitations, lab-in-field experiments have been adapted to terrorism prevention and violent extremism evaluations and studies. These experimental studies incorporate behavioral science “games” or scenarios focused on eliciting information on individual behaviors and prosocial actions, sometimes involving monetary or material incentives. Lab-in-field experiments can range in scope and focus, and can involve testing behavioral proclivities (e.g., resource or information sharing among otherwise at-odds or conflicting populations). For more information, see Cameron Sumpter, “Lab-in-Field Experiments for the Reintegration of Violent Extremists: The Promise of Prosocial Evaluation,” Washington, DC: RESOLVE Network, 2021.
53. For example, see Analytic Services Inc., “Using Gatekeeper Training as a CVE Tool”; and Weine et al., “Evaluating the Safe Spaces Program.”

About the Authors

Kateira Aryaeinejad specializes in research on violent extremism, focusing on research processes, global and local dynamics, and connecting rigorous research to policy and practice. Her work has focused on topics ranging from the evolution of and local support for violent extremist groups and narratives, online and offline communications, and peacebuilding and conflict dynamics. She holds an M.A. in International Peace and Conflict Resolution with a concentration in International Conflict Management from American University's School of International Service.

Thomas Leo Scherer is the research director for fp21. His work is focused on identifying and conducting high-quality research to improve U.S. foreign policy. He has previously worked on violent conflict research at the Center for Peace and Security Studies at the University of California San Diego and at the United States Institute of Peace. He has a Ph.D. in International Relations from Princeton University.

