Assessing the Impact of the Boston CVE Pilot Program: A Developmental Evaluation Approach

By Elena Savoa, Megan McBride, Jessica Stern, Max Su, Nigel Harriman, Ajmal Aziz and Richard Legault
Assessing the Impact of the Boston CVE Pilot Program: A Developmental Evaluation Approach
By Elena Savoa, Megan McBride, Jessica Stern, Max Su, Nigel Harriman, Ajmal Aziz and Richard Legault

Abstract
This article describes the use of developmental evaluation as applied to countering violent extremism (CVE) programs. It discusses the application of this method to an evaluation of the Boston CVE Pilot Program, with specific attention given to two CVE initiatives that were awarded pilot grants and volunteered to be evaluated. Developmental evaluation is an inherently iterative engagement that requires a continuous exchange between the parties designing the initiative and the parties evaluating it (the evaluation, in short, is not merely a post hoc engagement). We present the results of the evaluations using qualitative and quantitative data. The conclusion points to lessons learned in the application of a developmental evaluation framework: an assessment of the viability, utility, and benefits of utilizing such an approach to assess the impact of CVE programs (versus a traditional evaluation rubric). It also discusses the limitations that an outside organization engaged in this evaluative work might face.

Keywords: countering violent extremism; program evaluation; terrorism; tolerance; prejudice

Suggested Citation

Introduction
In 2016, Boston’s Executive Office of Health and Human Services (EOHHS) began the process of identifying and funding local efforts to combat violent extremism. Over the next year, this led to the creation of two different programs – an initiative to teach students to “say no to hatred and prejudice” and an initiative to support the integration of Somali immigrants in the United States by providing guidance in career development and job searches. Research suggests that there are over 1,400 countering violence extremism (CVE) programs around the world. Missing, though, is data on the impact of these programs. As one study noted, there is a problematic “dearth of publicly available data” on the evaluations of CVE efforts.¹ A RAND report published in 2017, found that only 95 CVE interventions included an evaluation component.² Of these, only one evaluation – at a university in Pakistan – used a rigorous methodological approach (the random assignment of participants to either a treatment or control group); and only one assessment – of a program in Montgomery County, Maryland – was of a CVE program in the United States.³ In other cases where such evaluations exist, they are often carried out by the same agency providing the service, introducing concerns about conflict of interest and the qualifications of those performing the evaluations. As RAND concludes, however, rigorous evaluations are required to inform decisions about whether to sustain, discontinue, or scale up these efforts.⁴
The reality, though, is that the evaluation of CVE programs is complicated by a series of political, conceptual, and logistical challenges. The label “CVE” has from its inception been politically controversial. Some critics asserted that CVE initiatives were stigmatizing Muslim communities and ineffective if not counterproductive. Others argued that the Obama Administration’s attempts to establish “community partnerships” (without specifying which communities) was too politically correct to be effective, and insufficiently focused on ideology. Others rejected CVE programs due to concerns that such initiatives were shaped by anti-Muslim bias and targeted already vulnerable populations. At its most extreme, these political issues resulted in some organizations actually rejecting CVE-linked grants. In 2017, for example, a number of organizations cited such concerns when rejecting over $2 million in CVE funds. Importantly, these issues aren’t unique to the United States; the United Kingdom’s PREVENT strategy, as just one example, was criticized and politicized on similar grounds.

Additionally, it is difficult to demonstrate the effectiveness of a CVE program given that terrorist violence is so rare. The absence of a radicalized individual within the population targeted by a modest CVE initiative is hardly proof of its effectiveness. Similarly, it is difficult to quantify program results given that the pathways to radicalization are so variable. It is thus nearly impossible, for example, to demonstrate that a CVE program reducing one risk factor in a community (e.g., perceived disenfranchisement) caused the lack of radicalization in that same community as there are far too many confounding variables. And finally, there are major logistical impediments to success: funding cycles for CVE initiatives and evaluation programs are not always in sync; funding for evaluation is sometimes inadequate to support the type of data collection necessary to produce scientifically valid results; and evaluations over short periods of time cannot be used to support long term conclusions about radicalization within a community.

There is no simple way to overcome this constellation of obstacles, and universal approaches to CVE and CVE evaluation will remain difficult. In contrast to one-size-fits-all initiatives designed and funded at the federal level, the Boston CVE Pilot Program was designed and funded at the local level. Our interdisciplinary team – composed of professionals with expertise in terrorism, public health, and evaluation – agrees that the lack of evaluation studies leads to lack of data on what interventions do and do not work. We also believe, though, that the lack of such studies leads to a lack of data on the type of evaluative approach that is best suited for evaluating CVE programs. And given the constellation of problems cited above, we believe that developmental evaluation – a relatively new form of evaluation – is a potentially powerful tool for tackling this challenge.

Developmental evaluation combines the work of assessment with a simultaneous effort to support and inform innovation. Because design and evaluation take place simultaneously, new elements are added and tested in an on-going feedback cycle. Innovations may thus be kept or discarded – in an iterative process that facilitates trial, error, and adaptation – depending on whether they contribute to the emergence of the desired outcome. Developmental evaluation is thus particularly well-suited to the design of social programs that remain in an exploratory phase, where specific outcomes are not yet known or are only vaguely defined. It is, in other words, particularly well-suited to the work of evaluating emerging CVE programs where stakeholders of diverse backgrounds are contributing to the development of initiatives designed to tackle a complex social phenomenon with no clear single causal factor. Moreover,
this approach enables the evaluator and the project developer to work together in an iterative process to develop and evaluate a program. In short, it facilitates collaboration and cooperation between those with different types of expertise (i.e., those expert in the dynamics of the local community and those expert in the processes of radicalization). Applying this method offers a clear path towards improving current and future CVE programming while simultaneously generating objective data evaluating the efficacy of this programming.

This article begins with an overview of the Boston CVE Pilot Program, transitions to a short explanation of developmental evaluation, and outlines the evaluation of two initiatives that were funded within the Boston CVE Pilot Program (and mentioned at the beginning of this article). A conclusion focuses on lessons learned in this process: an assessment of the viability, utility, and benefits of evaluating CVE programs within a developmental evaluation rubric (versus a traditional evaluation rubric), and a discussion of the limitations that an outside organization engaged in this evaluative work might face.

The Boston CVE Pilot Program

In March 2014, the United States National Security Council (NSC) chose three pilot cities (Boston, Los Angeles and Minneapolis) to test policy initiatives aimed at preventing “violent extremists and their supporters from inspiring, radicalizing, financing, or recruiting individuals or groups in the United States to commit acts of violence.”11 This pilot program, a foray into the work of countering violent extremism, marked the start of CVE in the United States. The Department of Homeland Security (DHS) defines CVE as “efforts focused on preventing all forms of ideologically based extremist violence, to include prevention of successful recruitment into terrorist groups.”12 As a 2019 RAND report noted, CVE programs largely fall into three categories with the “early phase focusing broadly on vulnerable populations, the middle phase focusing more narrowly on individuals at risk of radicalizing to violence, and the late phase focusing on individuals who have broken the law and are already involved in the criminal justice system.”13

The CVE initiative in Boston was informed by the argument that involving members of the community can improve the design of CVE programs.14 As a result, with the support of the Department of Justice, the Federal Bureau of Investigation, DHS and the National Counterterrorism Center, and under the coordination of the U.S. Attorney’s Office for the District of Massachusetts, a range of stakeholders in the Greater Boston region met over the course of six months to develop “A Framework for Prevention and Intervention Strategies: Incorporating Violent Extremism into Violence Prevention Efforts” (hereafter the Boston Framework).15 These stakeholders included community representatives, faith-based leaders, educators, mental health experts, and local government officials attuned to the concerns of the local populations and communities; it also included academics and policymakers who were knowledgeable about the challenges of preventing violent extremism.
The Boston Framework, which was issued in February 2015, established a set of guiding principles to assist local communities aspiring to “build resilience and capacity to prevent individuals, including young people, from being inspired and recruited by violent extremists.”\(^\text{16}\)

The text did not include a definition of resilience, but the context suggests that the authors understood “resilience” to mean the capacity to defend successfully against attempts to inspire individuals in the community to commit violent acts. The text identified a number of issues associated with violent extremism for which it proposed generic solutions with the understanding that the details would be developed by individual communities within the Greater Boston Area based on “local conditions” (i.e., conditions unique not to Boston writ large, but to the specific community attempting to respond to this challenge).\(^\text{17}\)

In addition to empowering locals, recent work in CVE has become increasingly concerned with the “dearth of publicly available data.”\(^\text{18}\) In part to address this issue, DHS and the National Institute of Justice (NIJ) have begun sponsoring evaluations of CVE programs in recent years. As part of this initiative, on October 1, 2015, our team (after being selected by DHS Science and Technology Directorate in a competitive process) initiated an evaluation of the Boston CVE Pilot Program.

At the time our team began evaluating the program, no specific CVE initiative had yet been funded, and it was still unknown what sort of initiative would take place and how the program would evolve. Though the leadership of the Boston CVE Pilot Program was originally located in a law enforcement agency (the Office of the US Attorney for the District of Massachusetts), it transitioned to a public health agency (EOHHS) as our evaluation efforts began. EOHHS ultimately issued a formal request for proposals (RFP), received four applications from local organizations, and awarded three modest grants.\(^\text{19}\) Our team was thus involved in the evaluation of the Boston CVE Pilot Program at several levels. We had access to those who had contributed to the original Boston Framework, those who were leading the Boston Pilot CVE Program, and those who had received grants to conduct pilot CVE initiatives. As a result, we had the opportunity to evaluate these efforts at each stage. Because we were involved in the project from conceptualization to implementation, we recognized that developmental evaluation (an inherently iterative engagement that requires a continuous exchange between the parties designing the initiative and the parties evaluating it) was the best-suited means of evaluating these initiatives.

For a variety of reasons discussed below, the specific pilot program initiatives that we were able to evaluate were only tangentially related to the prevention of ideologically based extremist violence or the prevention of successful recruitment of individuals into terrorist groups. As a result, the research findings we present in this article may be of less interest than the lessons learned from our use of developmental evaluation.
Developmental Evaluation and the Boston CVE Pilot Program

What is Developmental Evaluation?
Traditional evaluation, the most widely-known type of evaluation, is most effective when the objective is to measure results based on pre-identified outcomes and metrics. It is thus particularly well-suited to “situations where the progression from problem to solution can be laid out in a relatively clear sequence of steps.” But traditional evaluation is poorly suited to assessing the impact of programs grappling with complex social situations. These situations may be complex for a variety of reasons laid out by Michael Quinn Patton: “how to achieve desired results is not known (high uncertainty), key stakeholders disagree about what to do and how to do it, and many factors are interacting in a dynamic environment undermining efforts at control, and making predictions and static models problematic.” This complexity necessitates an evaluative approach that is more adaptive and fluid. Developmental evaluation is an approach designed to function in this type of a dynamic environment. As noted in “DE 201: A Practitioner’s Guide to Developmental Evaluation,” developmental evaluation follows the following principles:

1. The primary focus is on adaptive learning rather than accountability to an external authority,
2. The purpose is to provide real-time feedback and generate learnings to inform development,
3. The evaluator is embedded in the initiative as a member of the team,
4. The [developmental evaluation] role extends well beyond data collection and analysis; the evaluator helps to inform decision-making and facilitate learning,
5. The evaluation is designed to capture system dynamics and [bring to the] surface innovative strategies and ideas,

The approach is flexible, with new measures and monitoring mechanisms evolving as the understanding of the situation deepens and the initiative’s goals emerge (adapted from Westley, Zimmerman & Patton, 2006).

In short, developmental evaluation is an approach that combines the work of assessment with a simultaneous effort to support and inform innovation. Because design and evaluation take place simultaneously, new elements are added and tested in an on-going feedback cycle. Innovations may thus be kept or discarded in an iterative process that facilitates trial, error, and adaptation; they may also be kept or discarded depending on whether they contribute to the emergence of the desired outcome. Developmental evaluation is thus particularly well-suited to the design of social programs that remain in an exploratory phase, where specific outcomes
are not yet known or are only vaguely defined. It is, in other words, particularly well-suited to the work of evaluating emerging CVE programs where stakeholders of diverse backgrounds are contributing to the development of initiatives designed to tackle a complex social phenomenon with no clear single causal factor. Applying this method, one in which evaluation and design are married in an iterative process that facilitates innovation, should offer opportunities to improve current and future CVE programming.

Developmental Evaluation of the Boston Pilot CVE Program

Consistent with developmental evaluation, the team’s early objectives were to gather opinions on the overall CVE initiative, identify recommendations for practice, and develop a logic model for the evaluation of specific interventions that might ultimately be funded.25

We began by identifying and reaching out to those individuals who had contributed to the Boston Framework and used convenience sampling (snowball technique) to identify a range of stakeholders who had experience in violence-prevention initiatives (i.e., we asked contributors to the Boston Framework to identify additional professionals and community leaders).26 Ultimately, we conducted 52 interviews with project stakeholders from 45 organizations. We broadly classified these stakeholders as personnel from community-based non-governmental organizations, government (including law enforcement and public-school personnel), academia, or healthcare organizations. In a series of semi-structured interviews with these individuals, we collected their input on a variety of issues including, but not limited to, how to solicit local organizations’ interest in developing CVE interventions and what such CVE interventions should entail. These interviews aimed to frame CVE and related prevention efforts in Boston. The methods, analysis, and results of these interviews are presented in detail in a publicly available report (“Evaluation of the Greater Boston Countering Violent Extremism (CVE) Pilot Program”) released in November 2016.27 In the context of this analysis, however, it is important to note that the project stakeholders in the Greater Boston area were greatly concerned about violence. The majority of interviewees pointed out that Boston is a violent city, but noted that the violence is largely confined to only a few neighborhoods and is perpetrated by a relatively small number of people who tend to be known to the police. Stakeholders attributed the violence to poverty and other social ills, such as failed housing policies, the widespread use of narcotics, and the availability of guns, problems that are concentrated in those few neighborhoods. Stakeholders viewed violent extremism primarily as an act of violence. For the stakeholders, violent extremists’ purported ideological motivation was of secondary importance.28 In fact, some spoke about the disproportionate attention paid to terrorism versus non-ideological gun violence even though that the latter results in significantly more deaths. Stakeholders concluded that violent extremism prevention programs to be implemented in the Greater Boston Area should adopt a comprehensive approach to the prevention of violence and not focus on any one form of violence. They warned that a narrow focus on ideology and/or extremism, rather than on the prevention of acts of violence regardless of motivation, would be counter-productive.
Developmental Evaluation of Community-Level Interventions

In August 2016 — after the results of the interviews were shared with project stakeholders, and after EOHHS received responses to its request for information – EOHHS issued an RFP inviting local organizations to develop CVE interventions in the Greater Boston Area. Four organizations responded, and three applicants were awarded funding (ranging from $45,000 to $105,000) for CVE interventions to be implemented within a timeframe of one year. The following two organizations volunteered to be evaluated by our team: Empower Peace (hereinafter referred to by the name of their project, Online4Good Academy) and the Somali Development Center (hereafter referred to by the name of their project, SAFE Initiative). In the following sections, we describe our evaluation of these two initiatives. All study procedures were reviewed and approved by the authors’ Institutional Review Board (IRB Protocol #IRB15-3748).

Evaluation of the Online4Good Academy

Online4Good intended to enroll students and teachers in a one-day Academy where they would learn to “say no to hatred and prejudice” through a program that would “culminate with the students formulating their own online social media campaign plans designed to promote tolerance and acceptance, in particular of cultural differences, in their schools and communities.” After the one-day event, some students decided to develop and implement a campaign in their schools. The following sections describe our approach and results in evaluating the impact of such campaigns.

Methodology

We began the evaluation of the Online4Good program by engaging in conversations with the award recipient (a tenet of developmental evaluation) and helping them to outline project goals. As part of this process, we provided a logic-model designed to help the awardee relate their initiative to potential outcomes. In a traditional evaluation, a logic model guides the collection of evidence so as to evaluate a program along a pre-determined trajectory. In a developmental evaluation, the logic model functions to guide development and evaluation; it is thus continuously modified based on the feedback received during the implementation of the initiative. Ultimately, the award recipient indicated a desire to increase acceptance of others among youth and articulated a belief that exposure to hate messages, experiencing discrimination, and negative attitudes towards other racial-ethnic groups could lower such acceptance and fuel violence. Academic research on this topic is admittedly conflicted (there is no consensus regarding a correlation between hate, violence, and discriminatory behavior) and this decision thus highlights both the realities of locally-developed CVE initiatives and the limits of even a developmental evaluation framework. The Online4Good initiative was proposed by local actors and was funded by EOHHS; while our team was able to help develop a program that could be evaluated using empirically sound survey tools (thus injecting a degree of academic expertise into the process) we were not in a position to fundamentally challenge or alter the goals of the program.
Outreach data of the initial one-day training event

Attendance records showed that approximately 100 students from 22 schools located in 21 towns participated in the one-day training. As shown in Table 1, the Online4Good Academy one-day training event reached a variety of schools, across geographic locations, socio-economic status, and racial-ethnic composition.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Frequencies/Descriptive statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geographic location</td>
<td>Northeast: 7 (31%)</td>
</tr>
<tr>
<td></td>
<td>Southeast: 6 (27%)</td>
</tr>
<tr>
<td></td>
<td>Western: 3 (14%)</td>
</tr>
<tr>
<td></td>
<td>Central: 3 (14%)</td>
</tr>
<tr>
<td></td>
<td>Boston: 3 (14%)</td>
</tr>
<tr>
<td>Type of school</td>
<td>High schools: 10 (46%)</td>
</tr>
<tr>
<td></td>
<td>Mixed high school and middle school: 4 (18%)</td>
</tr>
<tr>
<td></td>
<td>Middle schools: 8 (36%)</td>
</tr>
<tr>
<td></td>
<td>Public schools: 18 (82%)</td>
</tr>
<tr>
<td></td>
<td>Private schools: 4 (18%)</td>
</tr>
<tr>
<td>Median household income of the town where the school is located</td>
<td>Mean = $73,160 (SD=$38,451)</td>
</tr>
<tr>
<td></td>
<td>Median = $69,829 Range ($18,226 – $199,519)</td>
</tr>
<tr>
<td>School diversity score (the probability that two randomly selected kids from the school belong to two different races or ethnic groups)</td>
<td>Mean = 0.26 (SD=0.14)</td>
</tr>
<tr>
<td></td>
<td>Median = 0.19 (Range=0.12 – 0.58)</td>
</tr>
<tr>
<td>Percentage of student population being white</td>
<td>Mean = 76% (SD=28%)</td>
</tr>
<tr>
<td></td>
<td>Median = 87% (Range=2.5% - 98%)</td>
</tr>
</tbody>
</table>

Evaluation of the impact of the campaigns

Due to the complexity of middle-school survey procedures and the lack of validated questionnaires for this age group, we decided to focus our evaluation efforts on high school students only. We initiated contact with the high schools’ teachers and determined that five out of the ten high school teams that attended the one-day event had moved forward with the creation of a school campaign. Therefore, we concentrated our evaluation efforts on these five
schools. We implemented a cohort study design with a comparative control group and gathered baseline and post-initiative data through an online survey. We also conducted interviews with the teachers implementing the program.

**Baseline Data Analysis**
Our evaluation began with the work of establishing baseline data. We needed to understand whether the populations targeted by the Online4Good initiative had been exposed to hate messages, experienced discrimination, or had negative attitudes toward other racial-ethnic groups. Before engaging in the development of the campaign, in October 2017, a convenience sample of students attending the selected and control schools were invited to complete an online survey. The survey was designed to assess attitudes related to the ultimate outcome of the initiative: acceptance of others. The survey included questions assessing student attitudes regarding cultural intelligence, acceptance of cultural differences, ethnocultural empathic awareness, experience of discrimination, use of social media, and exposure to hate messages (see Table 2 for definitions and the Appendix 1 for a copy of the survey). Control schools were selected to match the initiative schools by the household income of the town where the school is located, and the school diversity score. Baseline survey data were gathered from 196 students, 37 of whom were engaged in the creation of the campaigns, 41 of whom were exposed to the campaigns but not engaged in their development, and 118 of whom attended the control schools.

**Baseline survey sample**
One hundred ninety six students completed the baseline survey. We computed descriptive statistics on socio-demographic characteristics, attitudes and characteristics of their schools. We also gathered data on student use of social media and exposure to hate messages, including where the message was encountered and to whom the message was targeted. See Table 2 for descriptive statistics. Our data reflect a convenience sample, which we remind the reader is not representative of all high school students in Massachusetts. However, this sample gave us the opportunity to investigate any associations between students’ reported attitudes and their characteristics (e.g., age, gender, etc.) as well as characteristics of the school they attend (e.g., school diversity) to derive findings that can inform the development of future initiatives in similar populations.

**Baseline survey instrument**
We conducted a literature review to identify the most appropriate survey instruments that could be used to measure acceptance of others as promoted by the Online4Good initiative. Based on this review, we identified the Cultural Intelligence Scale (CQS) and The Scale of Ethnocultural Empathy (SEE) as the most appropriate instruments to be used in the context of our evaluation. More specifically, we used the Motivational Cultural Intelligence CQS subscale, the Acceptance of Cultural Differences SEE subscale and the Ethnocultural Awareness SEE subscale. The chosen instruments have a history of being tested in young populations. After cognitive testing, we determined that a minor re-wording of the questions was necessary for our target audience. For the sake of simplicity, in the following text we will refer to these subscales as Motivational CI, Acceptance of CD, and EC Empathic Awareness. Factor analysis results supported the aggregation of items into scales. Cronbach alpha of the sub-scales were
0.8 or higher. Table 2 includes the descriptive statistics of these scales. We interpreted these scales as follows:

- “Motivational CI” is a measure of the extent of students’ active pursuit of learning about and functioning in multi-cultural environments
- “Acceptance of CD” refers to the passive acceptance, appreciation, and understanding of differing racial-ethnic cultural traditions
- “EC Empathic Awareness” refers to the acknowledgment of structural and cultural racism in society

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Overall sample</th>
</tr>
</thead>
</table>
| Age             | Mean=16 (SD=1.3)  
                 | Median=17 Range (14-19) |
| Gender          | 67%  
                 | 33% |
| Race            | 65%  
                 | 35% |
| Grade           | 25%  
                 | 12%  
                 | 17%  
                 | 46% |
| Academic performance (What have been most of your grades up to now at school?) | 37%  
                 | 48%  
                 | 15% |
| Median household income of the town where the school attended by the student is located. | Mean=$62,861 (SD $16,772)  
                 | Median=$67,807 Range=$45,893-$140,268 |
| Diversity of the school attended by the student (the probability that two randomly selected kids from the school belong to two different races or ethnic groups) | Percentage of students attending schools categorized by diversity score:  
                 | Diversity score < 0.2=58%  
                 | Diversity score ≥ 0.2=43% |
### Friends of different races
*(Do you have friends of different racial-ethnic background?)*

<table>
<thead>
<tr>
<th>None</th>
<th>Few (1-2)</th>
<th>Some (3-5)</th>
<th>Many (&gt;5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4%</td>
<td>18%</td>
<td>36%</td>
<td>42%</td>
</tr>
</tbody>
</table>

### Experienced discrimination due to race/ethnicity

15%

### Ethno-cultural Empathic Awareness

**Numerical scale:**

- Mean=12 (SD=4)
- Median=12.5 Range=0-16

**Binary variable:**

- Low awareness: (score < 13) 50%
- High awareness: (score ≥ 13) 50%

### Motivational Cultural Intelligence

**Numerical scale 0-20**

- Mean=15 (SD=4)
- Median=16 Range=0-20

**Ordinal variable:**

- Low motivational CQ: (score ≤13) 28%
- Medium motivational CQ: (13 < score ≤ 19) 48%
- High motivational CQ: (score >19) 24%

### Acceptance of Cultural Differences

**Numerical scale 0-20**

- Mean=18 (SD=2)
- Median=20 Range=7-20

**Binary variable:**

- Low acceptance: (score < 20) 41%
- High acceptance: (score ≥ 20) 59%

### Social Media Use and Exposure to Hate Messages

<table>
<thead>
<tr>
<th>Have Social Media Profile</th>
<th>94%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use Social Media Daily</td>
<td>84%</td>
</tr>
</tbody>
</table>

### Type of Social Media Used

<table>
<thead>
<tr>
<th>Social Media Used</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>YouTube</td>
<td>74%</td>
</tr>
<tr>
<td>Instagram</td>
<td>70%</td>
</tr>
<tr>
<td>Snapchat</td>
<td>69%</td>
</tr>
<tr>
<td>Twitter</td>
<td>47%</td>
</tr>
<tr>
<td>Facebook</td>
<td>36%</td>
</tr>
</tbody>
</table>

### Exposure to Hate Messages (Verbal or Written) in the Past Week

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>20%</td>
</tr>
<tr>
<td>Rarely/very rarely</td>
<td>39%</td>
</tr>
<tr>
<td>Occasionally</td>
<td>25%</td>
</tr>
<tr>
<td>Frequently/very frequently</td>
<td>16%</td>
</tr>
</tbody>
</table>
Among Those Exposed to Hate Messages

<table>
<thead>
<tr>
<th>...Encountered on</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Media</td>
<td>50%</td>
</tr>
<tr>
<td>Verbal Speech by a Person They Knew</td>
<td>26%</td>
</tr>
<tr>
<td>Verbal Speech by a Stranger</td>
<td>25%</td>
</tr>
<tr>
<td>TV</td>
<td>15%</td>
</tr>
<tr>
<td>Music</td>
<td>11%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>...Targeted Against</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>54%</td>
</tr>
<tr>
<td>Sexual Orientation</td>
<td>37%</td>
</tr>
<tr>
<td>Religion</td>
<td>33%</td>
</tr>
<tr>
<td>Gender Identity</td>
<td>22%</td>
</tr>
</tbody>
</table>

Results of the Evaluation of the Online4Good Academy

Statistical methods and results for the evaluation of the Online4Good Academy are provided in detail in Appendix 2. Below we present a discussion of the main findings.

How can baseline survey results inform future initiatives?

The results of the baseline survey allow us to describe a population of high school students targeted by the Online4Good initiative, as well as students from schools with similar socio-demographic characteristics (control schools). As noted above, this population is neither representative of high school students in Massachusetts nor of high school students in the United States. However, associations found between variables are still to be considered statistically valid within the studied sample.

One of the most interesting and useful findings was the fact that students who acknowledge that some ethnic and racial groups might experience institutional and cultural barriers in society (institutional and cultural racism) are more likely to be motivated to interact with peers from other cultures. These results are consistent with another study conducted in Utah.\(^{37}\) Not surprisingly, those who have experienced discrimination themselves (the majority of whom in our sample were non-white) seem to have more awareness of racism (meaning that non-white minorities were more likely to be motivated to interact with peers from other cultures). Fortunately, awareness of institutional barriers can be taught and educational programs can be designed to enhance knowledge about these issues. We believe social media is a powerful platform through which these educational programs could be implemented, as 94 percent of our respondents reported having a social media profile. Our results also indicate that having more than five friends belonging to different racial-ethnic groups is a strong predictor of Motivational CI. While the result may not be surprising, the simplicity of such an indicator may be of practical use to assess Motivational CI in a given population by the use of short surveys.
Did the Online4Good initiative work?

Our results show that the Online4Good initiative worked to improve student attitudes as measured by Motivational CI and Acceptance of CD. Interestingly, two phenomena were occurring simultaneously. The desired attitudes of students in the control group declined over the two-month period of observation, and the desired attitudes of students exposed to the initiative improved over the same time. In tandem, this clearly demonstrates the positive impact of the program on students attending the schools where the initiative took place. The students who engaged in the development of the campaigns had the best outcomes in terms of changes in Motivational CI and Acceptance of CD, followed by the target of their campaign (classmates), while students in the control group fared the worst. This was particularly true for boys compared with girls. However, we note that independent of exposure to the initiative (i.e., in both control and participating schools), having friends of multiple racial-ethnic groups (baseline survey) or having acquired an awareness of racism over the course of the project (measured as a change between the baseline and post surveys in ethnocultural empathic awareness) protected some students from a decline in the desired attitudes.

Limitations

We believe the strongest limitation of our study – outside the CVE-related issues cited above about program length, the difficulty of demonstrating that a CVE initiative has worked, and the contested relationship between hate messaging and violence – is the lack of availability of adequate survey instruments for the attitudes being measured. While we used previously validated instruments to measure Motivational CI and Acceptance of CD, such instruments had not been tested in such a young population prior to this study. We did not know if the questions asked were relevant to the current generation of youth for measuring attitudes such as cultural intelligence. Future research should focus on developing appropriate instruments for this age category. In addition, the measures of discrimination we used have been widely used for non-white populations and focus on experiences of discrimination, mainly on the basis of belonging to a racial or ethnic minority group, but they lack testing on the white population. Indeed, we defined “hate” but did not differentiate between hate directed to oneself versus others (secondhand hate exposure). In general, more research should be conducted to understand the meaning of “hate” for this generation and how it impacts their mental health. Our results are also based on a small sample, due to the pilot nature of the initiative, and this may have limited our ability to further study other meaningful relationships. An additional limitation of the study design is that all students from a particular school were either assigned to the Online4Good or the control intervention. This may have resulted in observed differences in Motivational CI changes between intervention groups when in fact the differences were a result of some other dissimilarities between the schools, and not the intervention. Likewise, the lack of observed differences in Acceptance of CD change between intervention groups may have been due to dissimilarities in the schools.
Evaluation of the SAFE Initiative

The Somali Development Center (SDC) is a small organization that provides daily support to Somali immigrants and a few other immigrant communities. The range of activities and services is wide and includes helping community members complete paperwork related to immigration requirements or access to benefits, navigating the job market, translating documents, providing psychological support to refugees, and aiding with daily family issues. The SAFE Initiative, an effort coordinated by the SDC, ultimately consisted of workshops focused on supporting the integration of Somali immigrants in the US by providing guidance in career development and job searches, with special outreach to women.

Methodology

As with the previous example, we began our evaluation by describing the types of initiatives proposed by the awardee, and the characteristics of the audience they hoped to reach. Our team then engaged in conversation with the award recipient and helped them to develop an initiative that addressed their core concerns, resonated with the existing literature, and lent itself to evaluation. The SAFE initiative developed in part because of the factors (i.e., integration of immigrants, career development, and outreach to women) that project implementers considered central to building community resilience towards violence and violent extremism. We quickly realized that it was not possible to assess the impact of a specific initiative, such as SAFE, as it could not be clearly distinguished from everything else happening at the center. All of the SDC’s activities are delivered in an interrelated manner in support of the overall mission of the organization to facilitate the integration of Somali community members in American society. However, the flexibility permitted by developmental evaluation made it possible to approach the SDC’s efforts more organically, and to instead study the community served by the SDC more broadly. As such, we chose to focus on factors we identified in conversation with SDC representatives, and that we believed aligned with both the Boston CVE Pilot Program’s objectives and the SAFE Initiative’s more modest goals. Specifically, we focused on documenting experiences of discrimination, exposure to hate messages, trust in various organizations and government agencies, and concerns about the well-being of their community with the intent of informing future initiatives rather than focusing only on SAFE.

Outreach data of the SDC

A critical first task in this effort was to monitor and assess outreach. Based on attendance records over a seven-month period, the SAFE Initiative reached approximately 100 people who attended six workshops and eight community meetings focused on workplace culture, economic self-sufficiency, and intergenerational cultural gaps. The majority of attendees were women.
Survey instrument

In addition to assessing outreach, we administered a survey to the clients of the SDC. The survey instrument included socio-demographic questions such as age, gender, years spent in the U.S., and employment status, as well as questions related to attendance at the SAFE workshops. There were also questions regarding trusted institutions and organizations, experience of discrimination, exposure to hate messages, and informal social control. Once again, we identified the Cultural Intelligence Scale (CQS) and The Scale of Ethnocultural Empathy (SEE) as the most appropriate instruments to be used in the context of our evaluation. And again, we used the Motivational Cultural Intelligence CQS subscale, the Acceptance of Cultural Differences SEE subscale and the Ethnocultural Awareness SEE subscale. Cognitive testing was implemented on nine individuals. Based on their feedback, revisions were made to the survey instrument prior to implementation.

Results of the SAFE Initiative

Statistical methods and results for the evaluation of the SAFE Initiative are provided in detail in Appendix 2. Below we present a discussion of the main findings.

Did can we learn from evaluating the SAFE Initiative?

We believe the most important piece of information that emerged from this survey is that the Somali population we surveyed suffers from a high level of discrimination, with nearly half the individuals surveyed reporting they had experienced some form of discrimination. The level of trust in government reported was particularly high in this sample (57 percent expressed high levels of trust), which is not consistent with results from national surveys on the U.S. population (twenty percent). This result may be due to desirability bias (i.e., respondents providing answers that they believe are desired by the person giving the survey). The decline in the level of trust by increased levels of education is also not consistent with results from polls of the general U.S. population. Interestingly, trust in institutions was associated with an increased willingness to interact with other cultures, which we interpreted as a willingness to integrate into broader American society. We found a negative association between Motivational CI and exposure to hate messages. The greater the exposure to such messages (which were reported to be more frequently experienced via television compared to other sources of communication), the lower the willingness to interact with other racial-ethnic groups. While our results do not prove the effectiveness of the SAFE initiative and are based on a cross-sectional study design where causality cannot be proven, we believe they can inform future programs to the extent that they suggest that initiatives are needed to reduce the discrimination faced by the Somali community, including firsthand and secondhand exposure to hate messages. Programs to reduce bias against Somali immigrants should be directed toward the professionals and citizens with whom the Somali community frequently interacts (e.g., neighboring communities, local hospitals, etc.)
Limitations

Similar to the limitations stated for the Online4Good initiative, we believe that the survey instruments presented a significant limitation as they were not designed for the Somali community. Future research should focus on developing measures that reflect such attitudes and are culturally appropriate. Also problematic was that the cross-sectional design adopted for this study did not allow us to monitor changes over time. However, it provided important information on factors potentially associated with community resilience and the impact of exposure to hate. As noted in the results, desirability bias was also identified as being of potential concern for the interpretation of data even if surveys were implemented by a member of the team of Somali origin.

Conclusion: Developmental Evaluation and CVE Initiatives

The Boston CVE Pilot Program, an initiative that consisted not only of the theoretical Boston Framework but also of the fully implemented CVE initiatives solicited by EOHHS, was an ambitious effort to empower local community actors. This approach was a response, in part, to a widespread concern that “failing to prioritize local needs risks reinforcing the marginalization that drives [violent extremism] and results in the development of programmatic approaches that are ineffective and unsustainable.” From its very inception, the Boston Framework solicited input from both community leaders and terrorism experts. As a result of our unique role in this iterative and expansive process, as an evaluative team with access to the initiative at each level, our conclusions are effectively tiered in that we are able to speak not only to the specific initiatives that were executed (see the discussion above) but also to the feasibility and usefulness of developmental evaluation as a means of assessing CVE initiatives.

Developmental evaluation is an evaluative approach that is specifically designed for initiatives that are formulated and executed in complex and dynamic environments. In an ideal situation, the developmental evaluation of a CVE initiative (1) facilitates innovation by incorporating and responding to feedback as the program takes shape and (2) ensures that project developers benefit from the expertise that lies with the evaluating team (assuming that members of the evaluating team have expertise in terrorism or CVE). This approach enables the evaluator and the project developer to work together in an iterative process to develop a program, the result of which can be measured and evaluated. We believe it is a particularly useful approach to adopt when a CVE initiative is still under development (i.e., when the program managers haven’t yet fully designed the program and/or where the environment in which the program is being executed doesn’t preclude mid-stream readjustment). In the best-case scenario, a developmental evaluation would increase the likelihood of success, facilitate innovation, ensure accountability, provide concrete information to funders, and contribute to the problematically scarce body of research on evaluated CVE programs. The reality, however, is that the best-case scenario rarely exists and our experience with the Boston CVE Pilot Program is particularly compelling insofar as it was meaningful and productive despite an imperfect evaluative environment.
One challenge was the short timeline. In an ideal situation, the programs and evaluations would have continued over several years, integrating the results of our team’s analysis in an iterative and adaptive process. This would have:

- Made possible an on-going relationship between evaluators and program managers – facilitating the iterative give-and-take central to developmental evaluation – that would have (ideally) resulted in continually improving programs; and

- Made possible the evaluation of the long-term outcomes of the CVE initiatives that were funded.

Violent extremism is a complex phenomenon and developing a program that demonstrably reduces its likelihood in a one-year funding cycle is nearly impossible. To begin, the variables being evaluated (such as increasing acceptance of others or community resilience) do not shift in a perfectly linear process, and even if the initiatives were successful, this success might not be evident in the months immediately following the program’s completion. Additionally, even if the programs evaluated – Online4Good and the SAFE Initiative – improved conditions linked to the reduction in violent extremism, it would be difficult to demonstrate a lasting impact on such a short timeline. Efforts to reduce the threat of terrorism, and to reduce violent extremism, require long-term investment and a long-term perspective; this is the work of running a marathon and not a sprint.

An additional challenge was the reality of local politics. As one report on the evaluation of CVE initiatives noted, the work is made difficult by “politically driven (and not evidence-based) P/CVE agendas combined with a sense of urgency around violent extremism/terrorism-related issues that results in unrealistically ambitious objectives and unclear program objectives and untested or overly ambitious theories of change.” Beyond forcing the articulation of unrealistic objectives, politics will also shape the very programs that are adopted and pursued. Boston was no exception. For reasons unclear to our team, funding was allocated under a CVE umbrella despite feedback from stakeholders that violence in general (and not violent extremism) was a more pressing concern in the Greater Boston Area. Local stakeholders may have strong opinions about the shape and structure of CVE initiatives and these may, at times, directly contradict opinions of the program funders, the consensus of scholarly research, or the data gathered during the evaluative process.

A further imperfection in the operating environment is the reality that local applicants for CVE grants will not likely be aware of the empirical literature on countering violent extremism. As a result, proposals submitted may pursue agendas best described as CVE-adjacent. An evaluation of a U.S. CVE initiative in New Jersey identified some programs as “CVE-relevant,” suggesting that the term captured initiatives that were “not necessarily labeled as CVE programming per se, but that [are] intended to produce outcomes that are theoretically linked to factors (reported in peer-reviewed literature) associated with preemption of violent extremism.” A CVE-adjacent initiative, by contrast, is one that may (or may not) be labeled as CVE programming, and that is intended to produce outcomes associated with the prevention of violent extremism, but that is not necessarily supported by peer-reviewed literature or
rigorous evaluation. Such programs may develop organically and intuitively (with input from local leaders), and will continue to be funded regardless of whether or not empirical evidence supports the work they propose (because local communities will move forward with programs that feel helpful, even if there is no evidence to suggest that they will help with the particular problem being targeted).

These challenges are unavoidable: funding cycles may remain short; political push-back will always exist; local programs will continue to be funded regardless of their agreement with the academic consensus. And yet despite these challenges, our work in Boston highlighted the usefulness of approaching evaluation from within a developmental model and yielded modest but fruitful results of the CVE programs that were funded.

- Our evaluation of the Online4Good Academy made clear that initiatives taken by schools to address hate messaging must include an educational component on EC Empathic Awareness (i.e., explaining and understanding structural racism in society). They also suggested that social media is a powerful platform through which educational programs could be implemented.

- Our evaluation of the SAFE Initiative suggests that reducing Somali immigrants’ exposure to discrimination is critical to the integration of Somali immigrants, as exposure to hate decreases motivation to interact outside of one’s racial-ethnic group.

In short, developmental evaluation – a flexible, iterative, and adaptive approach that informs program managers as the programs unfold – is a particularly powerful means of evaluating CVE initiatives. In the case of the Boston CVE Pilot Program, developmental evaluation was a critical tool in helping to elucidate and consolidate diverse stakeholders’ attitudes toward CVE; communicate these findings to those overseeing the implementation of the program; work with local actors as they developed targeted CVE programs; and evaluate the short-term impact of these programs.
Appendix 1: School Survey

Questions 1-3 – Name of school and student ID

4. Have you ever participated in a campaign or activity that says no to hatred and prejudice, and/or that promotes social good?

☐ No
☐ Yes. Please describe:

Questions 5 & 6 – Name of the program the students participated in and level of engagement

7. Select your grade

Questions 9 & 10 – Questions for teachers engaged in the initiative – level of supervision and subjects taught

11. What is your age?

12. What gender do you identify with?

☐ Male
☐ Female
☐ Rather Not Say
☐ Other. Please specify:
13. What race/ethnicity do you consider yourself? Please select as many as you see fit:

- American Indian or Alaska Native
- Arab
- African American
- Native Hawaiian or other Pacific Islander
- Non-Hispanic White
- Non-Hispanic Black
- Asian
- East Asian
- Central Asian
- Western Asian
- Southeast Asian
- South Asian
- Haitian
- Hispanic
- Somali
- Don’t know Rather not say
- Other (please specify)

14. Do you have friends of different racial/ethnic backgrounds?

- None
- Few (1-2)
- Some (3-5)
- Many (>5)

15. What is your present religion, if any?

- Christian (non-denominational)
- Protestant
- Roman Catholic
- Mormon
- Orthodox Christian
- Jewish
- Muslim
- Buddhist
- Hindu
- Atheist (do not believe in God)
- Agnostic (not sure if there is a God)
- Nothing in particular
- Spiritual
- Don’t know
- Rather not say
- Other (please specify)
16. Have any of the following scenarios happened to you before in which you felt YOU WERE BEING TREATED UNFAIRLY? (Please select all that apply)

- Watched closely or followed around by security guards or store clerks at a store or the mall
- Got poor or slow service at a restaurant or food store
- You were treated badly by a bus driver
- Got poor or slow service at a store
- You were treated unfairly by a police officer
- Accused of something you didn’t do at school
- Unfairly called down to the principal’s office
- Got grades you didn’t deserve
- Treated badly or unfairly by a teacher
- Watched more closely by security at school
- Someone didn’t want to be friends with you
- You had the feeling someone was afraid of you
- Someone called you an insulting name
- People hold their bags tight when you pass them
- Someone made a bad or insulting remark about your race, ethnicity, or language
- Someone didn’t want to play or hang out with you
- Someone was rude to you
- People assumed you were not smart or intelligent
- You didn’t get the respect you deserved
- You weren’t chosen for a sports team
- Teachers assumed you weren’t smart or intelligent
- You’re called on less than your peers in class by teachers
- Your parents or other family members were treated unfairly or badly because of the color of their skin, language, accent, or because they come from a different country or culture
- You were in a car with your family that was unfairly pulled over by police
- You were walking on the street and were stopped and questioned by police
- Your family was treated unfairly by U.S. Customs Officials when entering the country via air, land, or water (e.g. airports, land borders, or piers)
- None of the above
- Other scenario that made you felt you were discriminated against. Please specify
17. For the scenarios that you have experienced before, which one bothered you the most? (Please select one)

- [ ] Watched closely or followed around by security guards or store clerks at a store or the mall
- [ ] Got poor or slow service at a restaurant or food store
- [ ] You were treated badly by a bus driver
- [ ] Got poor or slow service at a store
- [ ] You were treated unfairly by a police officer
- [ ] Accused of something you didn’t do at school
- [ ] Unfairly called down to the principal’s office
- [ ] Got grades you didn’t deserve
- [ ] Treated badly or unfairly by a teacher
- [ ] Watched more closely by security at school
- [ ] Someone didn’t want to be friends with you
- [ ] You had the feeling someone was afraid of you
- [ ] Someone called you an insulting name
- [ ] People hold their bags tight when you pass them
- [ ] Someone made a bad or insulting remark about your race, ethnicity, or language
- [ ] Someone didn’t want to play or hang out with you
- [ ] Someone was rude to you
- [ ] People assumed you were not smart or intelligent
- [ ] You didn’t get the respect you deserved
- [ ] You weren’t chosen for a sports team
- [ ] Teachers assumed you weren’t smart or intelligent
- [ ] You’re called on less than your peers in class by teachers
- [ ] Your parents or other family members were treated unfairly or badly because of the color of their skin, language, accent, or because they come from a different country or culture
- [ ] You were in a car with your family that was unfairly pulled over by police
- [ ] You were walking on the street and were stopped and questioned by police
- [ ] Your family was treated unfairly by U.S. Customs Officials when entering the country via air, land, or water (e.g. airports, land borders, or piers)
- [ ] None of the above
- [ ] Other scenario that made you felt you were discriminated against. Please specify

18. About the scenario which bothered you the most, how often has this happened?

- [ ] Very Frequently
- [ ] Frequently
- [ ] Occasionally
- [ ] Rarely
- [ ] Very Rarely
19. About the scenario which bothered you the most, why do you think it happened?
Please select as many as you see fit.

- The color of my skin
- My race
- My ethnicity or culture
- My language
- My accent
- My age
- My sex/ gender
- The clothes I wear
- The music I listen to
- My sexual orientation
- Any other reason. Please describe

20. About the scenario which bothered you the most, how did it make you feel?
Please select as many as you see fit.

- Angry
- Mad
- Hurt
- Frustrated
- Sad
- Depressed
- Hopeless
- Powerless
- Ashamed
- Humiliated
- Strengthened
- Other (please specify)

21. About the scenario which bothered you the most, how did you deal with it?
Please select as many as you see fit.

- Ignored it
- Accepted it
- Spoke up
- Kept it to myself
- Lost interest in things
- Prayed
- Tried to change things. Please describe in the Comment Field below:

________________________________________________________________________
________________________________________________________________________
22. In the past seven days, how frequently did you come across verbal or written expressions against a specific group because of their race, religion, disability, sexual orientation, ethnicity, gender, or gender identity?

- Very frequently
- Frequently
- Occasionally
- Rarely
- Very rarely
- Never

23. Please specify which characteristic(s) the verbal or written expressions were targeting against. Please select as many as you see fit.

- Race
- Religion
- Disability
- Sexual orientation
- Ethnicity
- Gender
- Gender identity
- Other (please specify)

24. Where did you come across the hate message(s)? Please select as many as you see fit.

- Verbal speech from a stranger
- Verbal speech from a person I know
- Poster or flyer on a wall
- Graffiti
- Social media such as Facebook, Instagram, Pinterest, Snapchat, Twitter, etc.
- TV
- Radio
- Music
- Book, newspaper, or magazine
- Other (please specify)
25. Which of the following social media tools do you use? (Choose all that apply)

<table>
<thead>
<tr>
<th>Social Media Tool</th>
<th>How often do you use the social media tool?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Twitter</td>
<td>☐ All the time</td>
</tr>
<tr>
<td></td>
<td>☐ Several times a day</td>
</tr>
<tr>
<td></td>
<td>☐ 1-2 times per day</td>
</tr>
<tr>
<td></td>
<td>☐ Several times per week</td>
</tr>
<tr>
<td></td>
<td>☐ Several times per month</td>
</tr>
<tr>
<td></td>
<td>☐ Never</td>
</tr>
<tr>
<td>Facebook</td>
<td>☐ All the time</td>
</tr>
<tr>
<td></td>
<td>☐ Several times a day</td>
</tr>
<tr>
<td></td>
<td>☐ 1-2 times per day</td>
</tr>
<tr>
<td></td>
<td>☐ Several times per week</td>
</tr>
<tr>
<td></td>
<td>☐ Several times per month</td>
</tr>
<tr>
<td></td>
<td>☐ Never</td>
</tr>
<tr>
<td>Google+</td>
<td>☐ All the time</td>
</tr>
<tr>
<td></td>
<td>☐ Several times a day</td>
</tr>
<tr>
<td></td>
<td>☐ 1-2 times per day</td>
</tr>
<tr>
<td></td>
<td>☐ Several times per week</td>
</tr>
<tr>
<td></td>
<td>☐ Several times per month</td>
</tr>
<tr>
<td></td>
<td>☐ Never</td>
</tr>
<tr>
<td>LinkedIn</td>
<td>☐ All the time</td>
</tr>
<tr>
<td></td>
<td>☐ Several times a day</td>
</tr>
<tr>
<td></td>
<td>☐ 1-2 times per day</td>
</tr>
<tr>
<td></td>
<td>☐ Several times per week</td>
</tr>
<tr>
<td></td>
<td>☐ Several times per month</td>
</tr>
<tr>
<td></td>
<td>☐ Never</td>
</tr>
<tr>
<td>YouTube</td>
<td>☐ All the time</td>
</tr>
<tr>
<td></td>
<td>☐ Several times a day</td>
</tr>
<tr>
<td></td>
<td>☐ 1-2 times per day</td>
</tr>
<tr>
<td></td>
<td>☐ Several times per week</td>
</tr>
<tr>
<td></td>
<td>☐ Several times per month</td>
</tr>
<tr>
<td></td>
<td>☐ Never</td>
</tr>
<tr>
<td>Salesforce Chatter</td>
<td>☐ All the time</td>
</tr>
<tr>
<td></td>
<td>☐ Several times a day</td>
</tr>
<tr>
<td></td>
<td>☐ 1-2 times per day</td>
</tr>
<tr>
<td></td>
<td>☐ Several times per week</td>
</tr>
<tr>
<td></td>
<td>☐ Several times per month</td>
</tr>
<tr>
<td></td>
<td>☐ Never</td>
</tr>
<tr>
<td>Network</td>
<td>Frequency Options</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Skype</td>
<td>☐ All the time ☐ Several times a day ☐ 1-2 times per day ☐ Several times per week ☐ Several times per month ☐ Never</td>
</tr>
<tr>
<td>Tango</td>
<td>☐ All the time ☐ Several times a day ☐ 1-2 times per day ☐ Several times per week ☐ Several times per month ☐ Never</td>
</tr>
<tr>
<td>MySpace</td>
<td>☐ All the time ☐ Several times a day ☐ 1-2 times per day ☐ Several times per week ☐ Several times per month ☐ Never</td>
</tr>
<tr>
<td>Digg</td>
<td>☐ All the time ☐ Several times a day ☐ 1-2 times per day ☐ Several times per week ☐ Several times per month ☐ Never</td>
</tr>
<tr>
<td>Flickr</td>
<td>☐ All the time ☐ Several times a day ☐ 1-2 times per day ☐ Several times per week ☐ Several times per month ☐ Never</td>
</tr>
<tr>
<td>Reddit</td>
<td>☐ All the time ☐ Several times a day ☐ 1-2 times per day ☐ Several times per week ☐ Several times per month ☐ Never</td>
</tr>
<tr>
<td>Instagram</td>
<td>☐ All the time ☐ Several times a day ☐ 1-2 times per day ☐ Several times per week ☐ Several times per month ☐ Never</td>
</tr>
</tbody>
</table>
26. Do you currently have your own profile on a social networking site like Instagram, Pinterest, Snapchat, Facebook, Twitter, or else?

☐ Yes
☐ No

27. If yes, how often do you use your social networking account? (If not, please select “Never”)

☐ Daily
☐ Weekly
☐ Monthly
☐ Less than monthly
☐ Rarely
☐ Never

I use other social media tool(s). (Please specify)

[Blank space for response]
28. The following statements ask about your thoughts and feelings in a variety of situations. For each statement, indicate how well it describes you by choosing the appropriate number on a scale from 0 to 4:

<table>
<thead>
<tr>
<th>Statement</th>
<th>0 (does not describe me very well)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4 (describes me very well)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. I enjoy interacting with people from different cultures.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>b. I am confident that I can socialize with people from a culture that is unfamiliar to me.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>c. I am sure I can deal with the stress of adjusting to a culture that is new to me.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>d. I enjoy living in cultures that are unfamiliar to me.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>e. I am confident that I can get accustomed to shopping while in a different culture</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>
29. The following statements ask about your thoughts and feelings in a variety of situations. For each statement, indicate how well it describes you by choosing the appropriate number on a scale from 0 to 4:

<table>
<thead>
<tr>
<th></th>
<th>0 (does not describe me very well)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4 (describes me very well)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. I feel irritated when people of different racial or ethnic backgrounds speak their language around me.</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>b. I feel annoyed when people do not speak standard English.</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>c. I get impatient when communicating with people from other racial or ethnic backgrounds, regardless of how well they speak English.</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>d. I do not understand why people want to keep their racial or ethnic cultural traditions instead of trying to fit into the mainstream.</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>
30. The following statements ask about your thoughts and feelings in a variety of situations. For each statement, indicate how well it describes you by choosing the appropriate number on a scale from 0 to 4:

<table>
<thead>
<tr>
<th></th>
<th>0 (does not describe me very well)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4 (describes me very well)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. I am aware of how society treats racial or ethnic groups differently than my own.</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>b. I recognize that the media often portrays people based on their racial or ethnic stereotypes.</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>c. I can see how other racial or ethnic groups are systemically oppressed in our society.</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>
d. I am aware of institutional barriers (e.g., restricted opportunities for job promotion) that discriminate against racial or ethnic groups other than my own.

Appendix 2 – Statistical Methods and Results

Evaluation of the Online4Good Academy: Statistical Methods and Results

Methods for baseline survey data analysis

We used simple and multiple ordered logistic regression to study the association between student characteristics, the characteristics of the schools they attend, and their attitudes toward other cultural groups. The dependent variables consisted of: an ordinal variable describing levels of Motivational CI (low, medium and high) and a binary variable describing Acceptance of CD (definitions are provided in Table 2). Independent variables were chosen based on teachers’ opinions on what might affect acceptance of others (recorded during interviews) and on theoretical relevance. Independent variables included age, gender, race, grade, academic performance, median household income of the town where the school is located, school diversity, exposure to hate messages, having friends of different races, experience with discrimination, and EC Empathic Awareness (definitions are provided in Table 2). We fit univariate binomial and ordered logistic regression models using logistical procedure in SAS 9.3 to test for associations between the two dependent variables (Motivational-CI and Acceptance of CD) and each independent variable. Prior to applying the ordered logistic model, we confirmed the parallel regression assumption by means of the Brant test. Multiple logistic models were created by including age, gender, and independent variables that showed a statistically significant association with the dependent variables in the simple models.

Results of baseline data analysis: simple models

Being a white student (OR = 0.5, 95% CI 0.3-0.9) or attending a diverse school (OR = 0.5, 95% CI 0.3-1.0) was inversely associated with Motivational CI. On the contrary, having friends of other racial-ethnic groups was associated with greater levels of Motivational CI (OR = 2.6, 95% CI 1.5-4.6). Interestingly, school diversity did not influence the number of friends of different races students had; in both diverse and less diverse schools, approximately 30% of the students had five or more friends of different racial-ethnic groups. Self-reported exposure to discrimination due to race and/ or ethnicity (OR = 2.4, 95% CI 1.1-5.3) and to hate messages (OR = 1.3, 95% CI 1.0-1.8) were associated with higher levels of Motivational CI. Data showed that females had twice the odds of reporting a higher level of Acceptance of CD compared to males (OR = 2, 95%
CI 1.0-3.8). Finally, EC Empathic Awareness was positively associated with both Motivational CI (OR = 3.2, 95% CI 1.8-5.8) and Acceptance of CD (OR = 3.9, 95% CI 2.1-7.4. No significant association was found between levels of Motivational CI or Acceptance of CD and age, school grade, academic performance, or median household income of the town where the school is located. For details on the computed statistics, please see Table 3.

**Results of baseline data analysis: multiple predictor models**

Multiple predictor models were fit to Motivational CI and Acceptance of CD, which included all significant predictors from the simple models described above plus age, gender, and race. In the multiple model for Motivational CI, students with higher EC Empathic Awareness had 2.4 times the odds of reporting a higher level of Motivational CI [OR=2.4, 95% CI 1.1-5.6] compared to those with lower EC Empathic Awareness. There was also a significant interaction between gender and friends of different races. Male students with more than five friends of other racial-ethnic groups had 7.1 times the odds of reporting higher levels of Motivational CI [OR=7.1, 95% CI 1.7-29.9] compared to those who had five or less, but for the female students, the odds ratio was 1.3 [OR=1.3, 95% CI 0.5-3.2]. This means that for males, having a diverse racial-ethnic network of friends has greater impact on cultural intelligence compared to females. In the multiple predictor model for Acceptance of CD, gender (female versus male) [OR=2.2, 95% CI 1.0-4.5], age [OR=0.7, 95% CI 0.6-1.0], experience of discrimination [OR=4.1, 95% CI 1.2-13.5] and EC Empathic Awareness [OR=3.6, 95% CI 1.8-7.3] were all significant predictors. For details on the computed statistics, please see Table 3.

**Table 3:** Association between students’ characteristics, motivational cultural intelligence and acceptance of cultural differences for the overall baseline sample of 196 students (Ordered logistic regression).

<table>
<thead>
<tr>
<th>Students’ characteristics</th>
<th>Motivational Cultural Intelligence (n=196)</th>
<th>Acceptance of Cultural Differences (n=196)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simple models OR (95% C.I.)</td>
<td>Multiple model OR (95% C.I.)</td>
</tr>
<tr>
<td>Age</td>
<td>1 (0.8-1.2)</td>
<td>1 (0.7-1.3)</td>
</tr>
<tr>
<td>Gender</td>
<td>1.2 (0.7-2.3)</td>
<td>-</td>
</tr>
<tr>
<td>Race (white versus non-white)</td>
<td>0.5 (0.3-0.9)</td>
<td>0.7 (0.3-1.6)</td>
</tr>
<tr>
<td>Grade</td>
<td>1 (0.9-1.3)</td>
<td>-</td>
</tr>
<tr>
<td>Academic performance</td>
<td>0.9 (0.6-1.3)</td>
<td>-</td>
</tr>
<tr>
<td>Town median household income</td>
<td>0.9 (0.7-1.2)</td>
<td>-</td>
</tr>
</tbody>
</table>
Pre-Post Campaign Data Analysis

Methods for pre-post campaign data analysis
The post-campaign survey was administered two months after the student-run groups launched their campaign. The calendar date varied for each school (March–May 2018) as some campaigns took longer than others to get started. We assigned each student a randomly generated ID that enabled us to match the post with the pre campaign survey data at the individual level. Of the 37 students participating in the creation of the campaigns, 26 (70 percent) completed both surveys and so did 39 classmates who were exposed to the campaigns but not involved in their development. In the control schools, 96 (81 percent) out of 118 students completed both surveys. Students attending the control schools were shown an online educational video.

<table>
<thead>
<tr>
<th></th>
<th>School diversity (diversity score &gt; 0.2 versus less than 0.2)</th>
<th>Exposure to hate messages verbal or written (continuous).</th>
<th>Exposure to hate messages verbal or written (categorical).</th>
<th>Friends of different races (5 friends versus less than five)</th>
<th>Experience of discrimination due to race/ethnicity</th>
<th>Ethno-cultural Emphatic Awareness</th>
<th>Friends of different races-by-gender interaction</th>
<th>Males, 5 friends vs less than five</th>
<th>Females, 5 friends vs less than five</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.5 (0.3-1)</td>
<td>1.3 (1.0-1.8)</td>
<td>Never</td>
<td>2.6 (1.5-4.6)</td>
<td>2.4 (1.1-5.3)</td>
<td>3.2 (1.8-5.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.6 (0.3-1.3)</td>
<td>0.9 (0.6-1.3)</td>
<td>Rarely/Very rarely</td>
<td>-</td>
<td>1.1 (0.4-3.3)</td>
<td>2.4 (1.1-5.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.7 (0.4-1.5)</td>
<td>0.9 (0.7-1.3)</td>
<td>Occasionally</td>
<td>reference</td>
<td>3.4 (1.2-9.4)</td>
<td>3.9 (2.1-7.4)</td>
<td>7.1 (1.7-29.9)</td>
<td></td>
<td>1.3 (0.5-3.2)</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>Frequently/Very Frequently</td>
<td>-</td>
<td>4.1 (1.2-13.5)</td>
<td>3.6 (1.8-7.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>reference</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
on how to prepare for a snowstorm emergency in lieu of participating in, or being exposed to, the Online4Good initiative. We applied analysis of covariance (ANCOVA) models to study the association between individual and school characteristics, and changes in Motivational CI and Acceptance of CD (defined as a change in score from the pre-survey to the post-survey) in the sample of 161 paired matched data. ANCOVA models were fit using PROC GLM in SAS version 9.3.

Results of pre-post campaign data analysis: simple models
We found that classmates of students participating in the Online4Good Academy, toward whom the campaign was likely targeted, experienced an increase in Motivational CI. Interestingly, students in the control group, showed a decline in Motivational CI. The difference in LS mean change between the two groups was 1.8 points (p-value=0.009) in favor of the participants in the Online4Good Academy initiative. Students in less diverse schools showed better improvements in Motivational CI after the initiative, compared to those in more diverse schools, we computed a LS mean difference of 1.43 points (p-value=0.0251) in favor of less diverse schools. Positive changes in EC Empathic Awareness were associated with increases in Motivational CI. Each one-point change in EC Empathic Awareness from the baseline survey was positively associated with a 0.21 point increase in Motivational CI (p-value=0.0074). Both students with more or less than five friends of different racial-ethnic groups showed a decline in Motivational CI over time, however a LS mean difference of 1.06 points (p-value=0.0667) was registered in favor of those with more than five friends of different racial-ethnic groups. In the simple models, only gender and age were significant predictors of change in the Acceptance of CD scale. Males had a greater decline in Acceptance of CD than females with a LS mean difference of -1.4 (p-value=0.0032). For a detailed description of the results, please see Table 4.

Results of baseline data analysis: multiple predictor models
Our results indicate that the following factors were significant predictors of change in Motivational CI: a student’s Motivational CI score at baseline, having more than five friends of different racial-ethnic groups (LS mean difference of 1.4 points (p-value=0.0182)), changes in EC Empathic Awareness (for every point in the EC Empathic Awareness score gained, from the baseline survey to the follow-up survey, students gained 0.27 points (p-value=0.0008) in the Motivational CI score). Results also indicated that there are differences in the impact of the initiative by gender. We detected a LS mean difference in Motivational CI of 4.9 points (p-value=0.0017) in favor of male participants over controls and 3.0 points (p-value=0.0637) for their male classmates over controls. We detected a LS mean difference of 2.0 (p-value=0.0291), 2.7 (p-value=0.0012) and 1.9 (p-value=0.0078) for female participants, their classmates, and controls over the male control group, respectively. Also, male participants had significantly higher changes in Motivational CI than female controls (LS mean difference = 2.9 points, p-value=0.0445). Even though the school diversity score was significant in the simple analysis, it was not included in the multi-factor analysis because 45 of the students had missing values for this variable.

It is also worth noting that students with higher baseline Motivational CI score tended to have smaller increases or larger decreases in Motivational CI change score (beta=-0.4, p-value<0.0001) due to a large number of students with a pre-initiative score at or near the maximum value.
Regarding the Acceptance of CD, our results indicate that the Online4Good Academy initiative did not have a significant impact on this scale (p-value = 0.2422). Overall, in both initiative and control schools, students experienced a decline in Acceptance of CD over time. The decline was worse for males, and students with less than five friends of a different race. Independently of gender and number of friends of a different race, students aged 15 and 16 were the only ones to experience a modest increase in Acceptance of CD, compared to other ages. For a detailed description of the results, please see Table 4.

### Table 4: ANCOVA Analyses for Least Squares Mean Change of MCI and ACD Scores

<table>
<thead>
<tr>
<th>Students’ characteristics</th>
<th>Change in Motivational Cultural Intelligence (n=161)</th>
<th>Acceptance of Cultural Differences (n=161)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simple models</td>
<td>Multiple model</td>
</tr>
<tr>
<td><strong>Categorical Predictors</strong></td>
<td>F-value (p-value)</td>
<td>LSM±SE (p-value*)</td>
</tr>
<tr>
<td>Initiative</td>
<td>4.12 (0.0182)</td>
<td>†</td>
</tr>
<tr>
<td>Controls</td>
<td>-1.4 ± 0.3</td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>-0.1 ± 0.7</td>
<td></td>
</tr>
<tr>
<td>Classmates</td>
<td>0.4 ± 0.6</td>
<td></td>
</tr>
<tr>
<td>Participants vs Controls</td>
<td>1.3 ± 07 (0.0855)</td>
<td></td>
</tr>
<tr>
<td>Classmates vs Controls</td>
<td>1.8 ± 0.7 (0.009)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>2.68 (0.1037)</td>
<td>†</td>
</tr>
<tr>
<td>Male</td>
<td>-1.4 ± 0.5</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-0.4 ± 0.3</td>
<td></td>
</tr>
<tr>
<td>Male vs Female</td>
<td>-1.0 ± 0.6 (0.1037)</td>
<td></td>
</tr>
<tr>
<td>School diversity (diversity score &gt; 0.2 versus ≤ 0.2)</td>
<td>5.15 (0.0251)</td>
<td>†</td>
</tr>
<tr>
<td>≤0.2</td>
<td>-0.1 ± 0.4</td>
<td></td>
</tr>
<tr>
<td>&gt;0.2</td>
<td>-1.5 ± 0.5</td>
<td></td>
</tr>
<tr>
<td>≤0.2 vs &gt;0.2</td>
<td>1.4 ± 0.6 (0.0251)</td>
<td></td>
</tr>
<tr>
<td>Friends of different races (5+ friends versus &lt; 5)</td>
<td>3.41 (0.0667)</td>
<td>†</td>
</tr>
<tr>
<td>&lt;5</td>
<td>-1.2 ± 0.3</td>
<td></td>
</tr>
<tr>
<td>≥ 5</td>
<td>-0.1 ± 0.4</td>
<td></td>
</tr>
<tr>
<td>&lt;5 vs ≥ 5</td>
<td>-1.1 ± 0.6 (0.0667)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Assessing the Impact of the Boston CVE Pilot Program: A Developmental Evaluation Approach
By Elena Savoa, Megan McBride, Jessica Stern, Max Su, Nigel Harriman, Ajmal Aziz and Richard Legault

<table>
<thead>
<tr>
<th>Age (categorical)</th>
<th>F-value</th>
<th>Beta±SE</th>
<th>F-value</th>
<th>Beta±SE</th>
<th>F-value</th>
<th>Beta±SE</th>
<th>F-value</th>
<th>Beta±SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>2.90</td>
<td>0.4 ± 0.2</td>
<td>0.88</td>
<td>0.2 ± 0.2</td>
<td>0.02</td>
<td>0.0 ± 0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14, 17 &amp; 18 vs 15 &amp; 16</td>
<td>4.06</td>
<td>(0.0196)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Initiative-by-gender interaction</th>
<th>F-value</th>
<th>Beta±SE</th>
<th>F-value</th>
<th>Beta±SE</th>
<th>F-value</th>
<th>Beta±SE</th>
<th>F-value</th>
<th>Beta±SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control – male</td>
<td>4.9 ± 1.5</td>
<td>(0.0017)</td>
<td>3.0 ± 1.6</td>
<td>(0.0637)</td>
<td>1.9 ± 0.7</td>
<td>(0.0078)</td>
<td>2.0 ± 0.9</td>
<td>(0.0291)</td>
</tr>
<tr>
<td>Control – female</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants – male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants – female</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classmates – male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classmates – female</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants – male vs Controls – male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classmates – male vs Controls – male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control – female vs Control – male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants – female vs Controls – male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classmates – female vs Controls – male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants – male vs Controls – female</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Continuous Predictors</th>
<th>F-value</th>
<th>Beta±SE</th>
<th>F-value</th>
<th>Beta±SE</th>
<th>F-value</th>
<th>Beta±SE</th>
<th>F-value</th>
<th>Beta±SE</th>
<th>F-value</th>
<th>Beta±SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>2.90</td>
<td>0.4 ± 0.2</td>
<td>0.88</td>
<td>0.2 ± 0.2</td>
<td>0.02</td>
<td>0.0 ± 0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change in Ethno-cultural Emphatic Awareness</td>
<td>7.40</td>
<td>(0.0074)</td>
<td>0.21 ± 0.1</td>
<td>11.72</td>
<td>(0.0008)</td>
<td>0.27 ± 0.1</td>
<td>0.55</td>
<td>(0.4614)</td>
<td>0.1 ± 0.1</td>
<td>0.31</td>
</tr>
</tbody>
</table>

† Because the interaction between initiative and gender was significant see results for interaction below.
‡ School diversity score not included in multiple predictor model because 45 students had missing values.
* P-value shown when the least-squares mean (LSM) difference between two levels of a factor were calculated.

Interviews with the Teachers
In keeping with a developmental evaluation approach, interviews were conducted with relevant parties after the campaigns were completed. Data from 12 semi-structured interviews with high school and middle school teachers were gathered and analyzed using NVivo version 11. Of the interviewees, nine participated in the initiative and three were from control schools. The interviews focused on understanding the environment in which the Online4Good campaigns developed (as well as that of the control schools), challenges in implementation efforts, and recommendations for future initiatives.
Teachers expressed concerns regarding student social behaviors, particularly discriminatory attitudes. In particular, they expressed concern about the use of derogatory language about other ethnoreligious or racial groups by students’ family members. They said that students’ exposure to such language in their family environment may normalize discriminatory attitudes against other racial-ethnic groups. Continuous exposure to derogatory language in the family environment and elsewhere, the teachers asserted, might also have an adverse effect on students in terms of emotional and overall mental health. One teacher was particularly concerned by her students becoming desensitized to the violence they were exposed to: “The things that they’re exposed to... they’re becoming so desensitized to...things, like violence or things like that, that concerns me, that they don’t think that’s a big deal, whether it’s fighting or the language.” Another teacher reported having witnessed an increase in bad behaviors and derogatory language out of ignorance rather than hatred. When asked about how to best implement future initiatives and promote acceptance of others and/or anti-bullying programs, teachers emphasized the need to incorporate such programs into the school’s curriculum and into the organizational structure of existing educational activities. This could include situations where students are brought together for other reasons (i.e. homerooms, assembly, school-clubs, advisory time, etc.). Teachers expressed concern about the unsustainability of many initiatives that are being initiated but do not have continued support and lack of specific expertise during implementation efforts. They also strongly encouraged efforts that include peer-led activities as students tend to pay more attention to their peers rather than adults, and suggested such activities be targeted to the 9th and 10th graders as they perceived this group to be more vulnerable than older students.

Evaluation of the SAFE Initiative:
Statistical Methods and Results

We obtained data from ninety-one individuals. The majority of survey respondents were women (70 percent), and the average age was 50 (SD=16). Geographically, respondents came from 17 towns across Massachusetts with 33 percent living in Boston (where the SDC is located). Also, 56 percent completed high school (either in the US or abroad), 26 percent had a fulltime job, 61 percent spoke only Somali at home, and 75 percent had been in the US for at least 10 years. Half of respondents had been clients of the SDC for at least 13 years, and 20 percent reported having attended at least one of the SAFE workshops. Of the 13 who attended the workshops, approximately 80 percent expressed general satisfaction. Eighty percent reported that attending the workshops enhanced their knowledge of the topics being discussed, that they could see themselves applying the knowledge acquired in their daily life and would recommend the workshop to others.
Trust
The great majority of respondents reported having a fair or great amount of trust in various types of organizations and professionals serving their community. In terms of trusted organizations and leaders, the highest level of trust was reported for religious leaders (96 percent), local city officials and healthcare providers (96 percent), followed by teachers (94 percent), the police (90 percent), other members of their community (88 percent), and the federal government in Washington, DC (84 percent). When all trust questions were aggregated into a scale, levels of trust were lower for those with a higher level of education. For each increase in the level of education, we found respondents reported 0.7 times the odds of trusting the above-listed organizations (OR=0.7, 95% C.I. 0.6-0.9). Interestingly, trust in such organizations was associated with greater Motivational CI. For each increase in the level of trust, respondents reported 1.2 times the odds of being at a higher level of Motivational CI (OR=1.2, 95% C.I 1-1.4).

Community concerns and experience with discrimination
When asked to name three leading concerns about their community, respondents mentioned in order of importance: youth wellbeing, housing, and education. When asked about whether they had experienced discrimination, 56 percent of participants reported having experienced unfair treatment attributed to bias, of which 33 percent occurred occasionally, 18 percent frequently or very frequently, and the remainder rarely. Interviewees perceived that the clothes they were wearing were the most frequent cause of discrimination (most of the survey respondents were women who wore a hijab, covered their heads, or wore typical Somali dress). Other examples of perceived discrimination included feeling that someone was afraid of them, being unfairly treated by a neighbor, seeing insulting graffiti on their property, being called a terrorist, receiving death threats, and many other situations of serious, unfair, and discriminatory treatment due to their ethnicity.

Exposure to hate messages
Thirty-three percent of respondents reported that during the past seven days they had come across hate messages (secondhand exposure). The most frequently cited source of exposure to hate messages was the TV (29 percent), followed by social media (13 percent), and verbal abuse from a stranger (10 percent). Thirty percent reported to have been exposed to offensive comments about their ethnic group, and 24 percent had been subjected to ethnically-charged name calling. Respondents that had been exposed to hate messages had 0.8 the odds of reporting high levels of Motivational CI (OR=0.8 95% C.I. 0.6-1). There was a significant association between the subjects’ level of education and reports of being exposed to hate messages (OR=1.3, 95% C.I. 1-1.6). Exposure to hate messages was not associated with trust in government.

Informal social control
We asked a series of questions about informal social control. Interviewees reported the likelihood that they could count on friends and family support in the following circumstances: if they saw acts of bullying (67 percent), someone being a victim of violence by a family member (66 percent), children disrespecting adults (64 percent), children seen skipping school (63 percent), youth discouraged by life circumstances (54 percent), children hanging out with troublemakers (55 percent), or someone manifesting signs of mental illness (49 percent).
instances where issues arise outside of the interviewees’ homes, informal social control seemed to be weaker, with fewer numbers of respondents being able to count on friends and family support if: someone was a victim of discrimination (33 percent), a fight broke out in front of their house (29 percent), or someone was a victim of police violence (29 percent).

About the Authors

Elena Savoia

Dr. Savoia is a Senior Scientist in Biostatistics and medical doctor by training. During the past fifteen years she has conducted research and training activities at the Harvard T.H. Chan School of Public Health focused on public health emergency preparedness. She is the deputy director of the Emergency Preparedness Research, Evaluation & Practice (EPREP) Program for which she has been leading research and training projects for the past six years. She has devoted her professional life to the use of quantitative and qualitative methods to measure public health systems’ capabilities in response to large-scale emergencies and assess populations’ behaviors and reactions in response to a crisis. During the past three years she has expanded her field of research to the evaluation of programs aimed at countering violent extremism. Dr. Savoia’s portfolio of activities include projects sponsored by the Centers for Disease Control and Prevention, Department of Homeland Security, National Institute of Justice, World Health Organization, North Atlantic Treaty.

Jessica Stern

Dr. Stern is a Fellow at the Harvard T.H. Chan School of Public Health and professor at the Pardee School of Global Studies at Boston University. Dr. Stern’s research focuses on perpetrators of violence and the possible connections between trauma and terror. She has written on terrorist groups across religions and ideologies, among them neo-Nazis, Islamists, anarchists, and white supremacists. She has also written about counter-radicalization programs for both neo-Nazi and Islamist terrorists. She has been working with a team at Boston Children’s Hospital on the risk factors for violence among Somali-refugee youth. She has held fellowships awarded by the John Simon Guggenheim Foundation, the Erik Erikson Institute, and the MacArthur Foundation. She was a Council on Foreign Relations International Affairs Fellow, a National Fellow at Stanford University’s Hoover Institution, and a Fellow of the World Economic Forum.

Megan McBride

Dr. McBride is a Visiting Postdoctoral Fellow. She is also a Postdoctoral Fellow at the Center for Strategic Studies (within the Fletcher School) at Tufts University, and a Research Analyst with a DC-area non-profit research and analysis organization. Prior to this work she was a Postdoctoral Fellow in National Security Affairs at the U.S. Naval War College, and a Middle East intelligence analyst with the National Security Agency. She holds a Ph.D. in Religious Studies from Brown University, an M.A. in Government from Johns Hopkins University, an M.A. in Liberal Arts from
the Great Books program at St. John’s College, and a B.A. in Psychology from Drew University. Her areas of expertise include terrorism, radicalization, religious and ideological violence, and theory of religion.

Max Su

Dr. Su is a Research Associate with backgrounds in mathematics, biostatistics, patient-centered outcomes research and information technology. He received his ScB in mathematics from Brown University and his ScD in Biostatistics from Harvard. Dr. Su is serving as the senior statistician and information technologist for the study. He has over 10 years of experience working with the randomized clinical trials pooled database in outcomes research.

Nigel Harriman

Mr. Harriman is a Research Coordinator working with the EPREP Program. He graduated from Cornell University College of Agriculture and Life Sciences in 2016 and majored in Biology and Society and minored in Infectious Disease. He has assisted in coordinating the data collection and analysis for several projects under DHS, NIJ, NATO, and CDC funding.

Ajmal Aziz

Ajmal Aziz is a Branch Chief at the Department of Homeland Security (DHS), Science and Technology Directorate (S&T) with broad experience in developing policies and analyzing homeland security and defense programs. He currently manages a robust portfolio of research and development programs aimed at public safety and security tailored to national and international stakeholders. He provides expertise, analysis, and advice to resolve, implement, and manage science and technology policy issues. He currently leads all multilateral international engagements at the Directorate and has extensive experience in engaging international stakeholders (primarily within the FVEY’s) to improve international cooperation on homeland security matters while developing relationships to advance collaboration with foreign nations to impact the formulation and implementation of national security policy.

Richard Legault

Richard Legault, Ph.D. is the Senior Advisor for Social Sciences, Science & Technology Directorate, U.S. Department of Homeland Security. Dr. Legault currently leads several research and development portfolios in a variety of domains. Some current portfolios include research and development activities for Terrorism Prevention; Human Trafficking; Social, Behavioral, and Economic Science support for technology implementation, online influence campaigns, and resilience. Dr. Legault has extensive experience in the application and management of quantitative and qualitative research methods. He has performed research in quantitative analysis of survey data, firearms ownership, policy evaluation, data usage and measurement, organizational violence, terrorism, and violence reduction strategies. He has published a book, book chapters, and several articles on these topics in peer-reviewed journals, including The Journal of Quantitative Criminology, Criminology and Public Policy, Crime and Delinquency,
Assessing the Impact of the Boston CVE Pilot Program: A Developmental Evaluation Approach
By Elena Savoa, Megan McBride, Jessica Stern, Max Su, Nigel Harriman, Ajmal Aziz and Richard Legault


Acknowledgments

This project was funded by the U.S. Department of Homeland Security (DHS), Science and Technology Directorate (Cooperative Agreement Number: 2015-ST-108-FRG005 Evaluation of the Greater Boston Countering Violent Extremism Pilot Program). The content of this manuscript as well as the views and discussions expressed are solely those of the authors and do not necessarily represent the views of DHS nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. government. We would also like to acknowledge Kathleen Deloughery and Daniel Woods for their guidance during the course of the project, Joshua Bartholomew, Souleymane Konate, Neiha Lasharie, and Saynab Yusuf for supporting data collection and providing feedback during the development of the manuscript, and Marcia A. Testa for guidance during data analysis.

Bibliography


Endnotes


8. The research team members who interacted directly with the CVE program managers both had advanced degrees in Public Health: Elena Savoia (Principal Investigator; MD, MPH) and Leesa Lin (Senior Program Manager; MSPH). Research team members who conducted the statistical analysis were Nigel Harriman (Research Assistant) and Maxwell Su (Research Associate with a background in mathematics and biostatistics; ScD). The team members who analyzed and contextualized the results were terrorism scholars Jessica Stern (PhD) and Megan K McBride (PhD). Finally, two representatives from the Department of Homeland Security provided guidance and comments on the final version of the manuscript: Rick Legault (PhD) and Ajmal Aziz.


17. Ibid.


24. Importantly, this emphasis on innovation is one of the features that differentiates developmental evaluation from formative evaluation. As Michael Quinn Patton has argued, formative evaluation is most appropriate when the task to be evaluated is already informed by a baseline body of knowledge (e.g., a set of best practices) and the objective is to make improvements “at the margins.” By contrast, developmental evaluation is most appropriate when there is inadequate baseline knowledge and it isn’t entirely clear what the best practices might be. In this case, you are working “further out on the continuum of the degree of uncertainty.” See: Inspiring Communities, “Developmental evaluation as alternative to formative assessment,” YouTube video.


29. The goals of the RFP were 1) To prevent violence and help to prevent people from joining organizations that promote, plan or engage in “violence;” and 2) to promote resilience by strengthening protective factors. The PEACE Project’s definition of violence was “an act that violates state or federal law and causes physical harm to a person, or property.” For a more specific definition, see: Massachusetts EOHHS, “Grant Application for the Massachusetts Promoting Engagement, Acceptance and Community Empowerment (PEACE) Project,” 2016.


31. For the view that hate is correlated with discriminatory or violent behavior, see:

32. For the view that hate is not correlated with discriminatory or violent behavior, see:

33. A “hate message” in this context is defined as verbal or written expression against a specific group because of the group’s race, religion, disability, sexual orientation, ethnicity, gender, or gender identity.

34. School diversity score is defined as the probability that two randomly selected kids from the school belong to two different races or ethnic groups.


43. Ibid.


Copyright

Copyright © 2020 by the author(s). Homeland Security Affairs is an academic journal available free of charge to individuals and institutions. Because the purpose of this publication is the widest possible dissemination of knowledge, copies of this journal and the articles contained herein may be printed or downloaded and redistributed for personal, research or educational purposes free of charge and without permission. Any commercial use of Homeland Security Affairs or the articles published herein is expressly prohibited without the written consent of the copyright holder. The copyright of all articles published in Homeland Security Affairs rests with the author(s) of the article. Homeland Security Affairs is the online journal of the Naval Postgraduate School Center for Homeland Defense and Security (CHDS).

Cover photo by Michael Skok on Unsplash