Consolidated Overview on: Evaluation of secondary and tertiary level P/CVE programmes

Radicalisation Awareness Network

RAN

Authored by **Paulo Teixeira**, RAN External Expert

Consolidated Overview on: Evaluation of secondary and tertiary level P/CVE programmes

2021

1. Introduction

Nowadays, the importance of external evaluations in the field of P/CVE practice is widely acknowledged. With a variety of programmes and approaches in this field it is crucial to comprehend what works and what does not work and to fully understand how these interventions promote change and impact on the problems they aim to tackle. Quality assurance measures as well as evaluations are confronted with the fast-moving and complex nature of extremist scenes, deradicalisation dynamics, socio-political developments and changing funding structures. Approaches must do justice to the varying characteristics of target groups, as well as to the variety of different stakeholders involved in planning, designing and implementing P/CVE measures. Evaluations are, however, indispensable to make grounded statements on good practices and in order to strengthen professional practices across the whole programmes and projects lifecycles.

In the vast field of secondary and tertiary prevention (including funding institutions, government authorities, civil society organisations), perspectives on how successful P/CVE programmes should look like, may also vary greatly. It is therefore necessary to develop a more nuanced understanding of the complex nature of secondary and tertiary prevention programmes (also regarding difficult access to target groups and sensitive data) to determine what can and should be reliably assessed by evaluations and under which circumstances. Apart from possibly diverging perspectives and unclear terminology, the process of conducting external evaluations may interfere with the daily work of first line practitioners, e.g. due to short funding periods and limited time during their daily work, but also due the sensitive information and complicated information sharing protocols, which is inherent to P/CVE work. Furthermore, there are important ethical issues that should be taken into account when designing external evaluations.

While there is a noticeable increase in the number and quality of P/CVE evaluations, some interventions still lack any type of evaluation or have only internal self-evaluation as a form of assessment, rarely focusing on change or impact. In our experience, M&E efforts in P/CVE initiatives usually tend to focus heavily on activities and outputs, not on assessing outcomes and broader impact on trends toward radicalisation or violent extremist activity. Some of these difficulties in focusing on change (outcomes or impacts) derive from weak programming, planning shortcomings or in general of intervention designs with poorly defined results and objectives. It is often noticeable that programmes and projects are not based on an underlying Theory of Change. Central question around 'What is the change we want to promote?' or 'How do the project's activities lead to meaningful change?' often remain insufficiently answered. (Project) planning and evaluation are two sides of the same coin, a poorly designed programme and/or project make the job of an external evaluator much harder.

When designing evaluations for P/CVE initiatives we face both, **analytical** and **practical** challenges. One of the most central analytical challenges is the difficulty of attributing change directly to programming efforts when evaluating initiatives, programmes or projects. The effects of secondary and tertiary prevention interventions cannot be tested in isolated settings. This means that there is no possibility to set up control groups, as would be the case in other research endeavors. On the side of practical challenges, data availability and reliability are usually the most common obstacles for evaluating change (outcomes and impact). This paper will further explore these challenges in depth.

Key objectives

The key objectives of this paper therefore present themselves as threefold:

1. To discuss the key challenges that come with evaluating secondary and tertiary prevention programmes;

2. To provide examples of successful approaches in this field that appropriately respond to the complexity of the issue and evaluation challenges;

3. To identify lessons learned from successful approaches and create a set of recommendations to practitioners and evaluators on how best to overcome the key challenges.

2. Challenges for evaluating secondary and tertiary level P/CVE programmes

When thinking about the specific role that evaluation can and should have in promoting the success of P/CVE activities, especially in turbulent times and unstable contexts where new threats emerge, and the challenges evaluators and external evaluations face one can look at those from a programme/project lifecycle perspective.

Key Challenges for Evaluations and Evaluators

1. Programme/Project Design Challenges

2. Analytical Challenges

3. Practical Challenges

Evaluation Design Challenges

Comprehensive planning is needed to choose an evaluation approach that fits the subject and needs of a project/ measure at hand. To come to an appropriate evaluation design is not always easy, but an important prerequisite for a successful evaluation. Hence, this paper starts with taking a closer look at the **Design Challenges** that may arise at the beginning of an evaluation endeavor. This includes the difficulty to have a clear and shared understanding of the applied concepts and terms.

Selecting a fitting methodological approach and deciding on metrics and helpful indicators can be challenging. Especially in cooperative multi-agency settings, evaluator(s) may have the role of uniting potentially diverging understandings at the outset and of structuring the evaluation process according the interests of all relevant partners.[1] This puts into question of what *actually* constitutes impact but there is also a lack of clarity on the defining characteristics of radicalisation and violent extremism (VE). We know that P/CVE programmes are implemented across a wide variety of cultural, social, and political contexts and local definitions and understandings of concepts such as VE, violence, community, tolerance, and even peace require contextually informed lexicons to develop meaningful programmes and measure effectiveness and impact. Sometimes these issues are amplified by weak research, analysis and/or understanding the specific problems/issues a programme/project is focusing on.

---

[1] Nehlsen et al., Evident and Effective

Better problem understanding is key to improve interventions and supporting the planning of an intervention in general. It may seem trivial to state that a clear understanding of a problem's causes and consequences will allow for targeted objectives and facilitate the selection of fitting indicators and metrics to assess change. But in absence of a holistic understanding of the specific problems that a given intervention is trying to address it is impossible to have a relevant and clear definition of what success looks like.

Devising indicators or metrics to measure relevant outcomes is another design challenge because even if the success of an intervention has been correctly defined, sometimes there are no established indicators or metrics to measure those outcomes of interest. In fact, in present practices, metrics (if used at all) vary considerably across interventions and validated scales or proxy indicators are not widely used or even considered as "valid measurements". It seems evident that when designing programme objectives that are focused on which activities will be conducted or which deliverables will be produced within a project lifetime instead of the intended changes these activities are to bring about, evaluations will end up measuring effectiveness and impact and not on *actual* change outcomes.

This focus on change can lead to more meaningful evaluations, that contribute and be one key element to promote sustainable change. With that in mind, it is important to clearly define the evaluation purpose, a clear understanding of the objectives of the evaluation as well as the evaluation scope. Evaluation designs need to address whether the analysis is focusing on a particular project, a policy theme or strategy or a broader range of programming that collectively contribute to CVE activities (which would be considered a multidimensional evaluation).

If we want to strengthen evaluation practices, and better identify new threats in the P/CVE field we need to aim for stronger evaluation research designs. In fact, P/CVE evaluations are often characterised by weak research designs and most take a largely descriptive approach using a single type of data and limited collection methods.

Analytical Challenges

As stated above, evaluations of P/CVE initiatives also face **analytical challenges**. One of the main analytical challenge is the historical difficulty of attributing change directly to programming efforts when evaluating initiatives/programmes/projects. Efforts to establish robust causal claims run into two major obstacles:

1) the impossibility of "measuring a negative" or proving that violent activity or radicalisation would have occurred had there not been an intervention, as it is challenging to predict what would have happened if no intervention had occurred.,

2) the hard-to-reach, scarce and very diverse populations of interest.

Interventions aim to address a wide range of stakeholders (individuals and groups), from those effectively at risk of radicalisation to those whose extremist views have already led to violence or need support to renounce violence. As deradicalisation or distancing from (violent) extremism tend to be fragile and non-linear processes, it is not easy to actively involve these target groups in an evaluation. Building required trust to access target groups apart from the stakeholders that are already involved with these individuals (e.g. social workers, civil society organisations, prison and probation staff etc.) takes time and that is usually a resource that is not abundant in evaluation processes.

When trying to establish causal links, traditional evaluation design evaluators use experimental or quasi-experimental designs ([2]) but these target groups are difficult to identify and that makes it problematic or

---

([2])non-experimental designs may be applied as well

impossible to establish a control or comparison group to compare the effects of an intervention with a counterfactual scenario. Ethically, it would also be highly questionable to withhold deradicalisation or prevention services from people who are considered in need of interventions/ prevention, only to be able to use them as a control group. This applies even more to the areas of secondary and tertiary prevention. To not provide distancing or deradicalisation support to individuals who are willing to leave extremist environments in order to conduct causality analysis would be morally indefensible. Especially, given the emerging security risks this would cause.[3]

This challenge is increased by the scope of time it can take from the end of an intervention until when certain outcomes and/or impacts start to be noticeable, establishing causality and isolating other factors (attribution) is very difficult. The longer the lag, the harder to have any kind of robust causality link. Secondary and tertiary prevention work can only be standardised and manualised to a very limited extent. In many cases, their setting can hardly be structured to the last detail by practitioners. Their constellations are generally diverse and dynamic.

Because of this causal attribution difficulty, it can be hard or even impossible to implement a counterfactual analysis. Most of the times, evaluations therefore opt to only make contribution claims. This means abstaining from attributing a specific weight to an intervention on a certain observable change and focusing on getting a strong set of data that can support a strong contribution claim.

Another analytical challenge is that P/CVE interventions occur in dynamic, ever changing contexts and political agendas. These unstable conditions severely limit the opportunities to establish a feedback loop throughout the evaluation process. Policymakers can, sometimes, disregard evaluation evidence and focus on one specific, predetermined policy option, while not considering all possible intended and unintended consequences.

The contextual nature of the problem is in fact a huge challenge for evaluations. Due to the very localised and context-dependent nature of VE and radicalisation drivers, it is extremely difficult to apply interventions and evaluation strategies evenly across multiple environments. The need for locally relevant indicators can limit the comparability of programmes in different contexts and that makes it hard to reach conclusions about the applicability of certain P/CVE interventions to other target groups or contexts. On the other hand, the effort to develop a robust portfolio of relevant, valid, and rigorous indicators for a specific local context can be the igniter for efforts to develop indicators for different contexts and promote meaningful comparative analysis.

Practical Challenges

When looking at the numerous **practical challenges** evaluators face when implementing their planned evaluation designs, data availability and reliability are some of the most common obstacles when trying to assess change (both, at the outcome and impact levels).

Practical challenges that are vastly documented in relevant literature regarding the field work challenges evaluations face, include the following:

- Difficulties on actively involving (all) relevant stakeholders in the evaluation process, as practitioners tend to have very tight and busy schedules.

- For funders, evaluations are often a valuable source regarding decision-making on the further support or discontinuation of a programme, hence practitioners' organisations may feel reluctant to agree to be externally evaluated. This is increased by the fact that evaluators are often only able to look at a fraction

---

[3] Möller et al., Zur Evaluation von Praxisansätzen in der Extremismusprävention

of cases, organisations work with, while the evaluation results tend to be interpreted to give conclusions about the approach of an organization as a whole.

- Not having the adequate or promised resources (financial, material or human) needed to correctly implement the previously validated evaluation design.

- P/CVE initiatives deal with sensitive political issues and data, hence local populations, government officials or programme staff may be reluctant to make certain information accessible.

- Ethical and security challenges may emerge while attempting to access relevant data and/ or identifying appropriate control groups.

- Sometimes evaluators end up with unnecessary quantities of data because of poorly defined evaluation objectives. Non aggregated data may then create more diffusion than focus.

- Difficulties in constructing baselines against which to measure the outcomes and impacts of projects and programmes.

- Absence of reliable and up-to-date data (like official statistics on success/ recidivism rates) against which to triangulate the results of evaluations.

The key message is that the sensitive and security-relevant nature of many questions asked during evaluations of secondary and tertiary level P/CVE programmes in an effort to assess attitudes and support for VE can reduce the reliability of the gathered data. Also, it is important to stress again that the indicators that are developed to measure impact and change must reflect local lexicons and realities. If that objective is not achieved, the evaluation findings are at risk of being inaccurate or irrelevant in relation to the context.

Main Ethical Issues when Evaluating P/CVE Programmes

- Deradicalisation or distancing from violent extremism can be a vulnerable, non-linear process. Any factor that may disrupt the trusting relationship between client and P/CVE practitioners poses as risk to the deradicalisation process itself.

- Collecting and storing evaluation data can pose risks to programme staff, interviewees, partners and respondents.

- The number of participants in tertiary P/CVE programmes is often small, hence public evaluation reports may allow for individual trace backs. Evaluations must ensure their identities remain protected.

- Data collection that includes potentially traumatic or sensitive topics risks emotional or psychological harm to respondents.

Most of the newer, considered to be more rigorous tools associated with monitoring and evaluation (M&E) for P/CVE have been developed in academic environments. In order to use these tools, specialised and trained evaluators as well as practitioners that know and understand their use and value are needed. Only the practitioners' involvement in the evaluation process can lead to more and better evaluation use.

However, there are usually limited opportunities for cooperation between the evaluators and researchers who develop these tools and the practitioners. Much is often lost in the translation from theoretical contexts to the application of new and more robust evaluation techniques to field-based programmes. More specific training and technical discussions spaces should be created in order to develop a better evaluation ecosystem where this dialogue could nurture better practices and incremental developments in evaluation tools and methods.

With that being said, many guides, publications, and frameworks have been published in recent years to promote more rigorous assessment of P/CVE interventions. These are materials that provide useful information on common indicators, metrics, data collection tools, and methodologies for M&E in P/CVE and aim to address some of the challenges faced by evaluators, organisations and practitioners in the field.

Summary of key challenges

This is a list of the main specific challenges faced by evaluators when trying to implement evaluations of secondary and tertiary level P/CVE programmes:

- Developing a shared understanding of the interventions and evaluations, their objectives and what success means.

- Programme and Project design flaws (interventions not focusing on change).

- Temptation to focus the evaluations on activities and outputs as they are easier to measure instead of focusing on meaningful changes.

- Lack or absence of a robust research phase that hinders real problem(s) understanding.

- Difficulty to make solid causality claims due to the multidimensional, sensitive and complex issues that are addressed in P/CVE programmes and projects.

- Ever changing contexts and the need to choose locally relevant indicators and metrics.

- Difficulty to access data needed to respond to evaluation needs and objectives in the best possible way.

- Lack of needed resources, either financial, material or human resources, to successfully implement defined and validated evaluation plans.

Further reading on common evaluation challenges

1. Baruch, B., Ling, B., Warnes, R. and Hofman,J. (2018). Evaluation in an emerging field: Developing a measurement framework for the field of counter-violent-extremism. Evaluation, 24:4, 475–495

2. Beaghley, S., Helmus, T. C., Matthews, M., Ramchand, R., Stebbins, D., Kadlec, A., & Brown, M. A. (2017). Development and Pilot Test of the RAND Program Evaluation Toolkit for Countering Violent Extremism. Santa Monica, CA: RAND Corporation.

3. Development & Training Services, Inc. (2015) CVE Evaluation: Introduction and Tips for CVE Practitioners. Development & Training Services, Inc.

4. Van Hemert, D., van den Berg, H., van Vliet, T.,et all, (2014) Synthesis Report on the State-of-the-Art in Evaluating the Effectiveness of Counter-Violent Extremism Interventions. Impact Europe.

3. Responding to Challenges for better secondary and tertiary level P/CVE programme evaluations

After identifying the challenges posed when developing and implementing evaluations of secondary and tertiary level P/CVE programmes, the following chapter will look into methodological approaches, strategies and instruments that can minimise or support to surpass these issues.

Develop a shared understanding of the problems and concepts

A frequently expressed concern is the uncertainty toward the choice of criteria on the basis of which secondary and tertiary prevention programmes are being evaluated. In order to demystify evaluation standards, it is recommended to discuss indicators by involving all relevant stakeholders from the beginning. To jointly answer questions similar to the following:

- What do we understand by concept Y? How do we define success? What are the key issues the programme is trying to tackle?

- To what extent is parameter X suitable to make a specific statement?

- How likely is it that we will be able to record these parameters in the course of the evaluation?

Such discussions can help to keep expectations of the evaluation realistic and harmonise potentially diverging interests among a group of different actors. This should be complimented by a comprehensive research or preliminary study phase by the evaluation team that adds to the discussion about first hand experiences with the programme's target groups. On the basis of jointly defined benchmarks, the review of the achievement of objectives can be worked out in advance or at the beginning of an evaluation. This process of joint clarification of expectations regarding the achievement of goals is relevant in every form of cooperation and in itself represents a useful outcome of an evaluation process. In the ongoing or completed course of action, reference to and review of these goals can thus take place and, for example, trigger reflections on the optimisation of processes

Also, when thinking about evaluation this will help to better define our evaluation scope, goals and the portfolio of questions the evaluation should answer and after that to better choose indicators and metrics.

Focus on Changes

To guarantee that we don't design evaluations focusing on what is done but on meaningful changes we should start by clarifying the programe or project evaluation logic and value chain. One specific kind of tool that can be useful to this is developing a **Logic Model** or a **Theory of Change** (ToC) of a P/CVE intervention that can help us to clarify the short, medium and long term changes associated with a certain programme or project. ([4]).

A Theory of Change (ToC) is an instrument that can be key in creating a shared understanding of the logic and objectives of a project or programme and, in doing so, is also helpful to define which questions should

---

([4]) Another important strategy is getting information on how the initiatives are perceived by the intended target groups (changes felt, initiatives value for these groups). This is important as these are the changes felt by target groups that are not always the expected outcomes defined in the planning phase of the projects or programmes.

the evaluation answer and define its objective(s). You should identify what you want to know and why you want to know it from the outset and a ToC is a great way to do just that.

The basis of a Theory of Change is to state the ultimate aim and impact you want to achieve and then describe what would need to happen to reach that point. The main elements of the theory of change model are:

- stating a clear aim or ultimate impact you want to achieve

- mapping the activities that we should undertake to achieve that programme/project aim

- mapping of how you will achieve your outputs and then the outcomes that work towards the ultimate aim

- understanding different ways of achieving the change you want

- build the evaluation model you want to implement

- identify the assumptions on which the programme/project was/will be built on ([5])

- map the external factors that could affect our effectiveness and we should take into account for developing mitigating strategies (in the planning phase) or to better explore root causes for problems or shortcomings during the evaluation process

This focus on change and the use of instruments like the ToC or Logic Models, can lead to more meaningful evaluations, that can be one key element to promote sustainable change. With that in mind, it's important and good practice to clearly define the Evaluation Purpose, a clear understanding of the objectives of the evaluation and the purpose/use for which it is being undertaken and the Evaluation Scope, as evaluations need to be clear about if their analysis is focusing on a particular project, a policy theme or strategy or a broader range of programming that collectively contributes to CVE activities (a multidimensional evaluation).

The Theory of Change (ToC) elements
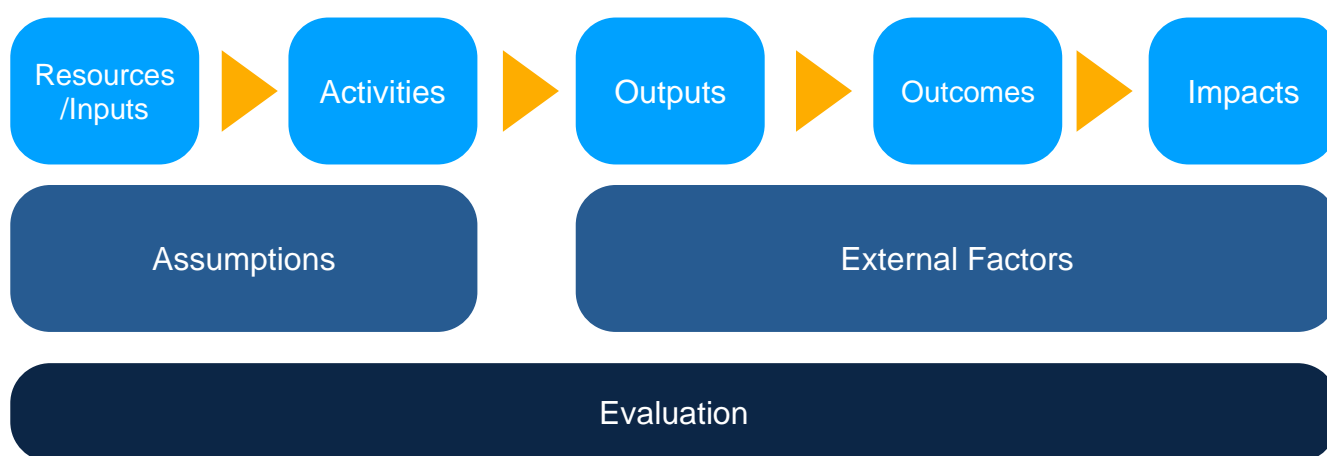
ToC elements at a glance:

- Resources include the human, financial, organisational and community resources a programme or project has available to direct towards doing the work. Sometimes this component is referred to as Inputs

---

([5]) The ToC is a tool that is useful both for planning a programme or project but also to plan the evaluation for a programme or project

- Activities are what the programme/project will do with these resources. Activities are the processes, tools, events, and actions that are an intentional part of the implementation process. These interventions are used to bring about the intended changes or results. Your Intended Results include all of the programme or project planned desired results (outputs, outcomes and impact).

- Outputs are the direct products/results of programme/project activities and may include types, levels and targets of services to be delivered by the programme.

- Outcomes are the short-term or medium-term changes relating any given activity outputs.

- Impacts are the fundamental intended or unintended changes resulting from the outcomes that occurs as a result of programme/project within a specific time scale.

- Assumptions are the belief system that we use to define the strategy of our programme/project.

- External factors are variables that cannot be controlled but can influence the real ability for a P/CVE programme or project to reach its goals.

In fact, a Theory of Change clearly expresses the relationships between actions and desired results and could also be described as a roadmap of the strategies and belief systems (e.g., assumptions, 'best practices', experiences) that make positive change in the lives of individuals and or a specific community. A Theory of Change can be articulated as a visual diagram that depicts relationships between initiatives, strategies and intended outcomes and goals.

In the process of developing a ToC, project members could start by clearly defining the problem or issue in a very robust way and discussing shared assumptions in relation to that problem definition. This definition of the "problem" is also relevant to pinpoint the audiences or target groups (primary and secondary) of a project or programme.

Advantages of Developing a Theory of Change in Evaluation processes:

1. It encourages a prospective not just a retrospective evaluation approach. The framework allows to specify the prerequisites for change and the steps to achieve them.

2. It focuses on the contribution towards the achievement of different stages of a process, rather than a diffusion of attention.

3. A clear understanding of the logic of the P/CVE interventions for evaluations and serves as a strong foundation to develop better evaluation questions and focus the evaluation on programme/project value chain.

4. By using the actual activities and objectives drawn from programme and projects analysis and focusing on the activities that practitioners actually do, it allows to move away from 'evaluation speech' and onto the actual ground work. This mirrors the process interventions actually take, without imposing additional frameworks onto the evaluation process

5. It makes explicit the underlying assumptions and relationships behind activities, allows these to be tested and related to final goals the programmes or projects are aiming to achieve

Further reading

1. Funnell, S., Rogers, P. (2011) Purposeful Program Theory: Effective Use of Logic Models and Theories of Change. San Francisco: Wiley/Jossey-Bass.

2. Hivos (2015) Theory of Change Thinking in Practice: A stepwise approach. Wageningen: Hivos. https://www.openupcontracting.org/assets/2017/09/Hivos-ToC-guidelines-2015.pdf

3. Peersman, G., Rogers, P., Guijt, I., Hearn, S., Pasanen, T., and Buffardi, A. (2016) 'When and how to develop an impact-oriented monitoring and evaluation system'. A Methods Lab publication. London: Overseas Development Institute.

4. Rogers, P. (2014) Theory of Change. Methodological Briefs on Impact Evaluation, No. 2. Florence: UNICEF Office of Research.

Changes in P/CVE Programmes

In addressing evaluation challenges and ways to overcome them in the Evaluations of P/CVE initiatives, in order to fulfil their promise and maximise their usefulness in helping with emerging threats of recent years like digital transformation, rising violent right-wing extremism (VRWE), issues of mental health and other vulnerabilities that have emerged due to the Covid-19 pandemic, we should aim to access/measure different areas of change.

In these programmes and projects, maintaining the goal of having change focused evaluations, we end up, most of the time, trying to evaluate changes in **attitudes**, **behaviours** and **activities** and on **relationships and social networks**.

Measuring changes in attitudes is important in most if not all P/CVE interventions evaluations. To measure changes in social, political, and ideological beliefs held by individuals targeted by an intervention, specifically their attitudes toward the use of violence and their ideological leanings. Change is commonly assessed by measuring an individual's knowledge of VE, as well as their perception of it.

The solution for above-mentioned challenges for evaluations could lie in the use of strategies that increase anonymity and the perception of anonymity when asking questions around sensitive topics to enhance the confidentiality of responses. The use of these data collection techniques, random response experiments, use a set of techniques so that respondents can answer a question without survey administrators knowing their individual responses. An example of a very practical way to access changes in attitudes is the use of Endorsement Experiments. These involve measuring the support for specific policies in a control group and a "treatment group". We ask members of a control group about their support for specific policies, and the same for members of the treatment group but these are also told that certain policies are supported by militant groups or VE Organisations. A comparison of the results gives us the extent to which knowledge of support by militant groups or VEOs for a policy altered or influenced responses, thus serving as an indirect measure of support for, or attitudes toward, VE. These methods aim to increase confidence levels and alleviate respondent's concerns about providing sensitive and potentially dangerous information.

Again, these approaches work better when supported by a strong ToC, and rigorous research, like pointed out before. Developing a robust Theory of Change (ToC) is an important tool for evaluators in order to identify intervention objectives and associated metrics, because each P/CVE relevant measure is explicitly linked to how the intervention aims promote change, either in attitudes, behaviours or relationships. At this time it is important to point out that ToCs need to be tested and refined throughout ongoing evaluations and underpinned with findings from relevant research that is focused on the drivers of violent extremism. The weakness of this type of metrics can be the underlying assumption about the relationship between extremist beliefs and violent activity.

A second level of change to be accessed focuses on behaviours and activities, like changes in individual engagement with VE groups and activities (including consumption of VE propaganda and online participation) or the opposite, participation in nonviolent acts or activities promoting tolerance or democratic processes. Changes in behaviours can be measured by using a mix of surveys, interviews, case studies and anecdotal evidence, as well as by collecting data on incidents of violence and violent offenders. One way of doing that could be looking at recidivism rates (i.e., incidents of relapse into violent or criminal activity) of former offenders these are all standard approaches to assess P/CVE interventions. However, only focusing on recidivism rates will exclude other important factors that allow for more comprehensive understandings of distancing from extremism.

Another way to measure changes in behaviours and activities is the use of life stories to understand/illustrate these changes, e.g. by using storytelling supported in a set of biographical interviews, this approach can/should be supported by visual elements, like photography and video. Visual Storytelling can be very effective to evaluate interventions.

Finally, it can be important to measure the relationships and social networks in P/CVE evaluations, but this is rarely done because of ethical and technical/logistical difficulties. However, measuring levels of cohesion, integration, and engagement of individuals in a community gives us relevant insights. In its research endeavour, the project 'DISLEX 3D' focuses on including not only the self-observations of VETOs but also the dimensions of observation by the systemic environment as well as those of professional accompaniment by P/CVE.[6]

---

[6] Find more information on the modus | zad project 'DISLEX 3D' here : https://modus-zad.de/en/themen/forschung-und-evaluation/dislex-3d/

Dealing with Ethical Issues

To deal with the ethical challenges that evaluators have to face when evaluating secondary and tertiary level P/CVE programmes including the sensitivity of the topics at hand, the following norms are considered good practice in social science and should be taken into account when conducting evaluations:

- **Informed consent:** whenever possible, respondents in interviews, focus groups, surveys, observations and other qualitative methods should give informed consent prior to data gathering. In some instances, obtaining informed consent may be impossible, for example in ethnographic research, where first a relationship needs to be built, which could be made impossible if respondents are asked to sign forms. In such cases it is best to ask for informed consent after data gathering.

- **Inclusion of minors:** the inclusion of minors in social research rules varies across countries. With the General Data Protection Regulation (GDPR) now in force, the rules spelled out there in relation to minors (and personal and sensitive data more broadly) need to be followed by those who control and process the data.

- **Confidentiality of research data:** data that respondents have provided should be kept confidential and stored securely. Only with explicit consent from the respondents can data be shared with anyone other than the researchers gathering the data.

- **The amount of time and effort requested from respondents:** respondents should be informed about what is expected from them and it should be assessed beforehand what amount of time is reasonable to ask from respondents.

- **Minimise harm and maximise good:** evaluations should be designed, conducted and disseminated with the aim of benefiting public good. Evaluators should give careful thought to the potential outcomes of evaluation and how findings might be used.

- **Emotional burden for respondents:** some questions might be a cause of distress for respondents, for example questions relating to painful incidents such as the experience of discrimination or failure. Questions related to mental health can also be uncomfortable to answer. Evaluators should pay attention to signs of stress and moderate or stop their questioning.[7]

Dealing with causality

As discussed before one of the hardest challenges for evaluators of secondary and tertiary level P/CVE programmes is to develop an evaluation design that could lead to a robust causality claim about the programme's impact on the problems it is trying to address.

Several different strategies may be used to undertake causal attribution, each of which has its own particular strengths, limitations and suitability according to the specific programme and evaluation context. While there are different ways of classifying designs and methods for causal attribution in evaluation, for the purposes of providing an overview in this document, there are three possible approaches presented below:

- **Counterfactual approaches** - aim to develop an estimate of what would have happened in the absence of a programme or project (or even policy) and compare this to what has been observed in the presence of the intervention. This approach requires that a control group or comparison group is used.

---

[7] Hofman, J. & Sutherland, A. eds. (2018), Evaluating interventions that prevent or counter violent extremism: A practical guide. Santa Monica, CA: RAND Corporation, https://www.rand.org/pubs/research_reports/RR2094.html.

- **Consistency of evidence with causal relationship** - this identifies specific patterns that would be consistent with a causal relationship, which is usually grounded in a well developed theory of change, and then seeks confirming and disconfirming evidence.
- **Ruling out alternatives** - this identifies possible alternative causal explanations and then aim to collect information to see if these can be ruled out or not.

A counterfactual approach involves developing an estimate of what would have happened in the absence of a programme or project, and comparing this to what has been observed in the presence of the intervention. Four types of evaluation design apply what we define as a counterfactual approach and these vary according on how the 'counterfactual '(non-intervention) effect is estimated.

Like stated previously, identifying a robust counterfactual for P/CVE interventions is challenging. When evaluating P/ CVE programmes, the ideal outcome involves identifying and proving a counterfactual or 'measuring a negative'. This would demonstrate that violence or radicalisation would have otherwise occurred had there not been an intervention. This challenge is most evident in prevention programmes and projects, where there is no guarantee that individuals would have become violent without intervention.

Table 1: Counterfactual Approaches

Counterfactual Approach	
Experimental Design	These involve the randomised assignment of participants to test the effects of an intervention. Participants are randomly allocated either to a group that participates in the programme/project, or to a control group that receives no intervention. These evaluations are called randomized controlled trials (RCTs). RCTs reduce the risk of selection bias, where participants and non-participants may have systematic differences other than whether they receive the intervention or not and because of that the apparent impact of the programme may be due to these differences rather than (or in addition) to its actual effect.
Quasi-Experimental Designs	These involve constructing comparison groups in various ways that do not involve random assignment. These designs are often more feasible in an evaluation of P/CVE programmes and may be considered to provide a sufficiently valid comparison between those who do and those who do not receive an intervention. There are several options for constructing comparison groups.
Hypothetical counterfactual	In some contexts/programmes it is possible to construct a 'hypothetical counterfactual' of what would have happened in the absence of the programme or policy by demonstrating that conditions would have remained the same.
Modelling	Involves the use of a statistically created counterfactual by developing a statistical model, such as a regression analysis, to estimate what would have happened in the absence of an intervention.

It should be pointed out that the use of experimental designs in the evaluation of secondary and tertiary level P/CVE programmes is usually not possible - there are ethical, practical, and security challenges in trying to identify control groups for P/CVE interventions. That is why when using a counterfactual approach to causality analysis there may need to be an increase of the use of quasi-experimental designs, when possible.

In other fields, the development and testing of several promising approaches for identifying and protecting control groups can be observed in different intervention areas. Approaches like the "stepped wedge" or "dynamic wait list" design where all individuals start in the "control group" and then sub-groups are randomly selected to undergo an intervention and this process is repeated until all individuals are in the "intervention

group" or the "switching groups" design where control and intervention groups are switched during the evaluation process.

When trying to use the '**Consistency of evidence with causal relationship'** approach to causal attribution, this means identifying what evidence would be consistent with a causal relationship, and then gathering and analysing data from different sources to determine whether the evidence matches this. This approach is normally guided by a theory of change (ToC), whether this is elaborated on in detail or implicit in the programme or project logic model of the intervention.

Different data collection and analysis methods can be used to assemble, but it can be advisable to combine several of these methods within a single evaluation according to the level of certainty required and the possible counter-explanations identified. In addition, evidence from previous research and evaluations can be used, as previous gathered data and evaluations could already prove certain causal links.

The below-assembled list includes some briefly identify some of the different data collection and analysis elements that can be used in this causality approach:

- Achievement of intermediate outcomes - checking whether all cases that achieved the final impacts also achieved the intermediate outcomes identified in the P/CVE programme ToC.

- Checking results against expert predictions - involves predictions based on the ToC or an emerging theory of wider contributors to the outcomes, and then following up on whether or not these predictions actually materialises over time.

- Checking timing of impacts - determining whether the timing of impacts is consistent with a causal relationship, again with reference to the ToC.

- Comparative case studies - systematically comparing case studies to understand the diverse factors that may be responsible for the programme outcomes and/or impacts.

- Checking consistency with existing literature - checking results against what is known from reviewing the literature in the area of P/CVE, to identify consistencies/inconsistencies.

- Interviewing key informants - not asking if they believe the intervention has produced the impacts it is instead about asking them to explain the causal processes following their involvement. The objective is to decode the perceived changes promoted by the P/CVE programmes in stakeholders.

- Modus operandi - using previous experiences of participants and stakeholders to determine what array or pattern of effects is typical for an intervention.

- Process tracing - developing alternative hypotheses and then gathering evidence (clues) to determine whether or not these are compatible with the available hypotheses.

- Qualitative comparative analysis (QCA) - we compare the configurations of different case studies to identify the key components that appear to be most responsible for producing specific outcomes.

- Realist analysis (Realist Evaluation) - using a realist theory of change (i.e., what works for whom in what circumstances and through what causal mechanisms) to identify contexts where a specific mechanism promotes positive results.

The third option identified for establishing causal attribution is to identify **possible alternative explanations** for the achievement of results, outcomes or impacts and then gather data to see if these can be ruled out. This is particularly useful to use when the available evidence may be sufficient only to suggest "correlation" but not "causality".

Options to operationalise this strategy include:

- Key informant interviews - asking experts in secondary and tertiary level P/CVE programmes, community members or other stakeholders to identify other possible explanations for changes and, if possible, to assess whether these explanations can be ruled out.

- Process tracing - using evidence to rule out alternative explanatory variables at each step of the P/CVE programme's ToC.

- Ruling out technical explanations - identifying and exploring possible ways that the observed results might reflect technical limitations of the gathered data rather than causal relationships.

- Modelling - investigating alternative explanations using statistical analysis such as regression or logistic regression to control for confounding factors.

- General elimination methodology - this is a process implemented in two stages, (1) identifying possible explanations (including that the observed changes are indeed due to the intervention, plus as many alternative explanations as possible) using a combination of options such as those identified above and (2) collecting and analysing data to verify if the possible alternative explanations can be ruled out.
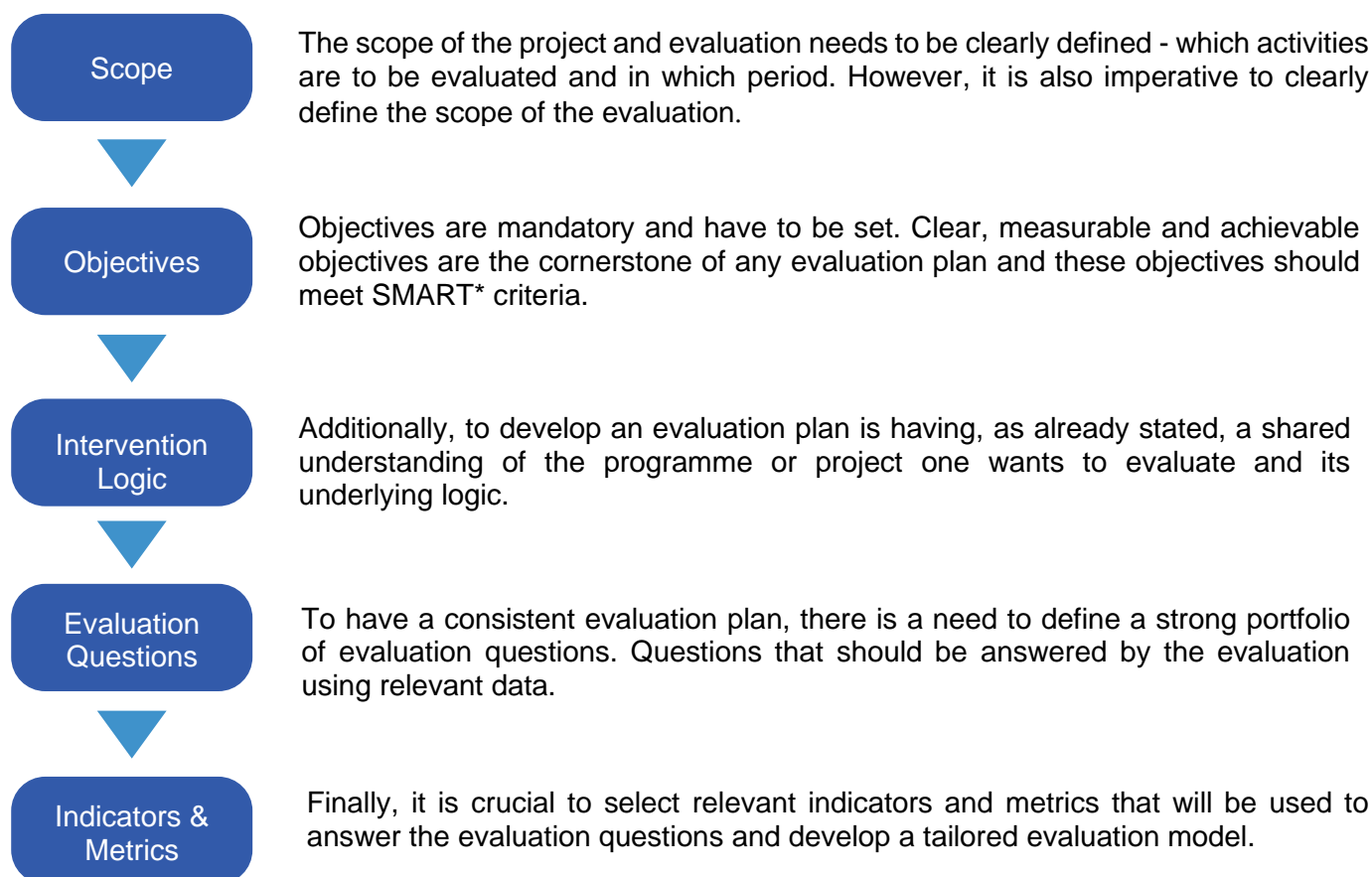
With these different methodological approaches available, it depends on specific implementation contexts, specific problems the programmes aim to tackle, access to target groups, resources (such as time and budget), evaluators skills and knowledge and specific evaluation's objectives, the final selection of the causal analysis to be used by evaluators.

Further reading

1. Braddock, K., (2020), Experimentation & quasi-experimentation in countering violent extremism: Directions of future inquiry. Washington DC: Resolve Network.

2. Peersman, G., Guijt, I., Pasanen, T. (2015) Evaluability assessment for impact evaluation. Guidance, checklist and decision support for those conducting the assessment. A Methods Lab publication. London: Overseas Develop

3. Phillips, C., & Goodman, K. (2006). Causal criteria and counterfactuals; nothing more (or less) than scientific common sense. Emerg Themes Epidemiol, 3.

4. Ris, L. and Ernstorfer, A. (2017). Borrowing a wheel: Applying existing design, monitoring, and evaluation strategies to emerging programming approaches to prevent and counter violent extremism. Briefing Paper, Peacebuilding Evaluation Consortium.

5. Rogers, P. (2014) Overview of Impact Evaluation. Methodological Briefs on Impact Evaluation, No 1. Florence: UNICEF Office of Research.

6. Rogers, P. (2014) Overview: Strategies for causal attribution. Methodological Briefs on Impact Evaluation, No 6. Florence: UNICEF Office of Research.

7. Rogers, P., Peersman, G. (2014) Addressing complexity in evaluation. Canberra: DFAT workshop.

8. Stern, E. (2015) Impact Evaluation: A guide for commissioners and managers. London: Bond.

4. Developing an Evaluation Plan

For effective evaluation of secondary and tertiary level P/CVE programmes, evaluators can choose different specific approaches and methods but there are always several steps one should follow in the development of a robust evaluation plan.

Scope — The scope of the project and evaluation needs to be clearly defined - which activities are to be evaluated and in which period. However, it is also imperative to clearly define the scope of the evaluation.

Objectives — Objectives are mandatory and have to be set. Clear, measurable and achievable objectives are the cornerstone of any evaluation plan and these objectives should meet SMART* criteria.

Intervention Logic — Additionally, to develop an evaluation plan is having, as already stated, a shared understanding of the programme or project one wants to evaluate and its underlying logic.

Evaluation Questions — To have a consistent evaluation plan, there is a need to define a strong portfolio of evaluation questions. Questions that should be answered by the evaluation using relevant data.

Indicators & Metrics — Finally, it is crucial to select relevant indicators and metrics that will be used to answer the evaluation questions and develop a tailored evaluation model.

*__SMART__: __S__pecific, __M__easurable, __A__ttainable, __R__esult-focused and __T__ime-specific

It is important to stress a strong participation from different stakeholders in all these phases is crucial. Higher participation levels will lead to a stronger more meaningful evaluation and also promote evaluation results' use.

Further, it is important to create a model of logical relationship from evaluation objectives, to evaluation questions, metrics, indicators (qualitative and quantitative), data sources and the information gathering instruments. In selecting indicators and metrics, evaluators should avoid the temptation of supporting effectiveness or impact claims in very limited sets of key performance indicators (KPIs) as it is not consistent with the problem's complexity. This is important as all the evaluation conclusions are to be supported by data collected in the evaluation process.

Objectives ▶ **Questions** ▶ **Metrics & Indicators** ▶ **Data Sources** ▶ **Data Collections methods**

It may be useful to cluster the evaluation questions in relevant categories such "efficiency", "effectiveness" or "impact". These categories can be important to help evaluators develop a strong evaluation narrative when analysing data as well as for the evaluation report itself and drawing of conclusions and recommendations.

Evaluation reporting should also be done in engaging ways, like making use of storytelling strategies, especially visual storytelling using photography and/or video. Further, a storytelling approach to data presentation using infographics and creative ways to visually present the indicators and metrics, can also be useful.

Finally, it is important to enunciate the importance of the evaluation team to be in close contact and work with the programme/project team during all phases of evaluation, from design to implementation as well as reporting.

Key Lessons

1. Validation of all phases of the evaluation plan development process with key stakeholders. Main evaluation questions, indicators and metrics, evaluation plan, relevant data and main conclusions should be viewed from a multi-stakeholder lens.

2. It is not helpful to over-claim or trying to hyper-simplify what is complex in nature. To avoid trying to access value or meaningful change based on a small set of what you can consider key performance indicators (KPIs)

3. Do not try to force an attribution narrative if contribution is the only strong narrative that can be made.

4. Draw conclusions based exclusively on the data that was collected within *this* evaluation process.

5. Make clear recommendations and explain why they are being made (always relating to your conclusions).

6. Create engaging ways to present findings to promote evaluation use and perceived added value

Recommendations

- More robust evaluation designs is a key area of investment when aiming for better evaluations. Better P/CVE evaluation designs usually are mixed method, use both quantitative and qualitative data, have a participatory dimension involving all relevant stakeholders and can have both a more formative and summative focus, according to their main goal and intended use but should definitively have a change/impact focus.

- Using a Theory of Change or a Logic Model is helpful throughout several points of a project's lifecycle. ToCs are useful in the planning stage to focus on change and in designing evaluations to pose better evaluation questions and choose meaningful indicators and metrics.

- Promote the use of data collection strategies that promote anonymity and its perception by key stakeholders. Data security is critical to ensuring confidentiality is maintained. The inherently sensitive nature of P/CVE programmes and projects means that participation in and the evaluation of P/CVE programmes can be extremely sensitive to the individuals involved.

- To mitigate the ethical issues present in secondary and tertiary level P/CVE programmes, evaluators must guarantee that all participants in a specific evaluation are exposed to the intervention so that they are all recipients of the potential benefits of that intervention. There are options that allow to overcome this ethical challenge, one is to use a switching-replications design, in which the initial control and treatment groups are switched during the evaluation process.

- The development of evaluation rubrics and choosing culturally, locally valued relevant indicators and metrics can also be very helpful to measure and interpret new threats of different "configurations" of traditional ones.

- Using participatory approaches is key, involving all relevant stakeholders in all phases of the evaluation, design, implementation, analysis and reporting leads to better design and data analysis but also to greater evaluation use.

- To maximise evaluations' usefulness – there lies a chance in visualisation strategies for gathering but especially for presenting information in more engaging ways. Visual storytelling, photographic and video based narratives or using stories told in the first person can lead to more change in perceptions and attitudes than traditional reporting.

About the author:

Paulo Teixeira is a Lisbon based sociologist with a Master in planning and evaluation of Development Processes. Paulo gives classes on planning and Evaluation at the Portuguese Catholic University and others. Finalising a PhD in Communication Studies, researching on the use of photography and visual storytelling in Communication for Social Change. Paulo is the owner of Logframe a consultancy firm, working in planning and evaluation of social Policies, Programmes and Projects, both in Portugal and Internationally. In 2016 Paulo launched Humane Focus, a humanitarian photography project and in 2018 created BUZZ!!!, a communication agency for the social economy sector. Paulo was a European Evaluation Society Board Member and is co-author of three books on Planning and Evaluation.

Bibliography

Baruch, B., Ling, B., Warnes, R. and Hofman,J. (2018). Evaluation in an emerging field: Developing a measurement framework for the field of counter-violent-extremism. Evaluation, 24:4, 475–495

Beaghley, S., Helmus, T. C., Matthews, M., Ramchand, R., Stebbins, D., Kadlec, A., & Brown, M. A. (2017). Development and Pilot Test of the RAND Program Evaluation Toolkit for Countering Violent Extremism. Santa Monica, CA: RAND Corporation.

Braddock, K., (2020), Experimentation & quasi-experimentation in countering violent extremism: Directions of future inquiry. Washington DC: Resolve Network.

Dawson, L., Edwards, C. and Jeffray, C. (2014), Learning and Adapting: The Use of Monitoring and Evaluation in CVE: A Handbook for Practitioners. London: Royal United Services Institute

Development & Training Services, Inc. (2015) CVE Evaluation: Introduction and Tips for CVE Practitioners. Development & Training Services, Inc.

Funnell, S., Rogers, P. (2011) Purposeful Program Theory: Effective Use of Logic Models and Theories of Change. San Francisco: Wiley/Jossey-Bass.

Hivos (2015) Theory of Change Thinking in Practice: A stepwise approach. Wageningen: Hivos. https://www.openupcontracting.org/assets/2017/09/Hivos-ToC-guidelines-2015.pdf

Hofman, J. & Sutherland, A. eds., (2018) Evaluating interventions that prevent or counter violent extremism: A practical guide. Santa Monica, CA: RAND Corporation. https://www.rand.org/pubs/research_reports/RR2094.html.

Oakden, J. (2013), Evaluation rubrics: how to ensure transparent and clear assessment that respects diverse lines of evidence. BetterEvaluation, Melbourne, Victoria, https://www.betterevaluation.org/sites/default/files/Evaluation%20rubrics.pdf

Patton, M. Q. (2010), Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use. New York: Guilford

Peersman, G., Guijt, P., Pasanen, T. (2015) Evaluability assessment for impact evaluation. Guidance, checklist and decision support for those conducting the assessment. A Methods Lab publication. London: Overseas Develop

Peersman, G., Rogers, P., Guijt, I., Hearn, S., Pasanen, T., and Buffardi, A. (2016) 'When and how to develop an impact-oriented monitoring and evaluation system'. A Methods Lab publication. London: Overseas Development Institute.

Phillips, C., & Goodman, K. (2006). Causal criteria and counterfactuals; nothing more (or less) than scientific common sense. Emerg Themes Epidemiol, 3.

Ris, L. and Ernstorfer, A. (2017). Borrowing a wheel: Applying existing design, monitoring, and evaluation strategies to emerging programming approaches to prevent and counter violent extremism. Briefing Paper, Peacebuilding Evaluation Consortium.

Rogers, P. (2014) Overview of Impact Evaluation. Methodological Briefs on Impact Evaluation, No 1. Florence: UNICEF Office of Research.

Rogers, P. (2014) Theory of Change. Methodological Briefs on Impact Evaluation, No. 2. Florence: UNICEF Office of Research.

Rogers, P. (2014) Overview: Strategies for causal attribution. Methodological Briefs on Impact Evaluation, No 6. Florence: UNICEF Office of Research.

Rogers, P., Peersman, G. (2014) Addressing complexity in evaluation. Canberra: DFAT workshop.

Stern, E. (2015) Impact Evaluation: A guide for commissioners and managers. London: Bond.

United States Agency for International Development (USAID), (2012) An Inventory and Review of Countering Violent Extremism and Insurgency Monitoring Systems. Prepared by Lynn Carter and Phyllis Dininio. Washington, DC: USAID.

Van Hemert, D., van den Berg, H., van Vliet, T.,et all, (2014) Synthesis Report on the State-of-the-Art in Evaluating the Effectiveness of Counter-Violent Extremism Interventions. Impact Europe.

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications from: https://op.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: http://eur-lex.europa.eu

Open data from the EU

The EU Open Data Portal (http://data.europa.eu/euodp/en) provides access to datasets from the EU. Data can be downloaded and reused for free, for both commercial and non-commercial purposes.