

Evaluating interventions that prevent or counter violent extremism

A practical guide

EDITORS

Joanna Hofman and Alex Sutherland

CONTRIBUTORS

Jacopo Bellasio, Megan Sim, Anke van Gorp, Kate Cox, Kei Ito, Ben Baruch, Benoit Guerin, Adam Martin, Sarah Grand-Clement, Barbara Janta, Catherine Saunders, Elta Smith, Gavin Cochrane and Marie-Louise Henham



For more information on this publication, visit www.rand.org/t/RR2094

Published by the RAND Corporation, Santa Monica, Calif., and Cambridge, UK

© Copyright 2018 RAND Corporation

RAND® is a registered trademark.

RAND Europe is a not-for-profit research organisation that helps to improve policy and decision making through research and analysis. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

Limited Print and Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law. This representation of RAND intellectual property is provided for noncommercial use only. Unauthorized posting of this publication online is prohibited. Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Permission is required from RAND to reproduce, or reuse in another form, any of its research documents for commercial use. For information on reprint and linking permissions, please visit www.rand.org/pubs/permissions.

Support RAND

Make a tax-deductible charitable contribution at
www.rand.org/giving/contribute

www.rand.org

www.randeurope.org

Preface

This document derives from the work carried out under the IMPACT Europe project¹ funded by the Seventh Framework Programme.² The project sought to identify what works in preventing and countering violent extremism (P/CVE) and focused on developing an interactive toolkit which helps evaluate interventions in this field.

RAND Europe led the work to develop the toolkit and identify evaluation methods and approaches that can be applied in evaluations

of P/CVE interventions to provide much-needed evidence on how to further improve these. This document synthesises the collection of methods and approaches included in the IMPACT Europe evaluation toolkit.³

This work was supported by the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 312235.

1 <http://impacteurope.eu/>

2 https://ec.europa.eu/research/fp7/index_en.cfm

3 <http://www.impact.itti.com.pl/index#/home>

Acknowledgements

We would like to thank all contributors for their time and inputs to this work. We are particularly grateful for the assistance of the members of the IMPACT Europe consortium for their invaluable inputs and suggestions to this collection during the development of the evaluation toolkit: Dr Dianne van Hemert and Dr Helma van den Berg (Nederlandse Organisatie voor toegepast natuurwetenschappelijk onderzoek), Dr Jean-Luc Marret (Fondation pour la Recherche Stratégique), Liliana Leone and Laura Tagle (Studio CEVAS), Dr Michał Tanaś (ITTI Sp. z o.o.), Prof Marcello Gallucci (University of Milan Bicocca), Dr Louis Logister (Hogeschool Utrecht), Dr Ron van Wonderen and Rozetta Meijer (Verwey-Jonker Instituut),

Declan Kirrane (Intelligence in Science), Ran Cohen (International Security and Counter-Terrorism Academy), Abdul-Rehman Malik (Radical Middle Way), Theo Muskee (Politie Nederland), Dr Harald Weilnböck (CULTURES INTERACTIVE e.V. Verein zur Interkulturellen Bildung und Gewaltprävention), and Mogens Schwensen (Sorø Kommune).

We also wish to thank Dr Megan Sim, Barbara Janta, Kei Ito and Talitha Dubow for their various contributions to different chapters. Finally, we would like to thank our quality assurance reviewers, Dr Sue Guthrie and Dr Christian van Stolk, for their constructive review and advice.

Table of contents

Preface	III
Acknowledgements	i
Table of contents	iii
Introduction	1
Developing this collection and user guide	9
Ethical considerations	17
Case studies	23
Comparisons/benchmarking	29
Contribution analysis	33
Cost-benefit analysis	37
Cost-effectiveness analysis	43
Cross-sectional study and data analysis	49
Data mining	53
Descriptive statistics	57
Desk-based research and literature review	63
Focus groups	69
Interviews	75
Logic models/theory of change	83
Longitudinal studies and data analysis	89
Meta-analysis	93
Network analysis	99
Objectives and options analysis	105
Observation techniques/ ethnography	111
Policy scientific approach	117

Qualitative data analysis and grounded theory	121
Quasi-experimental designs	127
Randomised control trials	131
Realist evaluation	139
Stakeholder analysis	143
Surveys	147
Appendix: Additional resources	153

CHAPTER ONE

Introduction

Joanna Hofman and Jacopo Bellasio

1.1. Current state of evaluation in the field of preventing and countering violent extremism

There has been wide acknowledgement of the need to prevent and counter violent extremism and considerable investment has been made in this area over recent years. However, these efforts have often been based on untested programmes and assumptions and have rarely been evaluated. Thus, it is unclear what works effectively and what does not when it comes to P/CVE interventions (Gielen 2017). Evaluation is an important tool that can help policymakers, programme implementers and other stakeholders learn whether their efforts bring the expected results and how to improve in future. Despite calls for a systematic investigation into successful and unsuccessful practices, lessons learned and an analysis of why certain procedures have or have not worked (European Parliament, 2017), there is still a paucity of relevant (and robust) evaluation literature.

One of the biggest efforts to date to map and systematically assess the state of P/

CVE evaluations concluded that most methodologies applied so far fell short of providing rigorous investigations of programme outcomes informed by empirical data (van Hemert et al. 2014; Feddes and Galucci, 2015). A quality assessment of evaluations was conducted based on scientific criteria taking into account the practical limitations that evaluators face when conducting field research. The methods and instruments commonly used in evaluation research in the area of behavioural sciences were taken as a reference point. In evaluation research, empirically testable assumptions and hypotheses should be outlined and tested. As such, the quality of evaluations was scored as follows (van Hemert et al. 2014):

- Low: no empirical investigation was conducted while circumstances would allow for a more thorough methodological assessment to answer key evaluation questions;
- Medium: empirical data were collected but the circumstances would allow for a more advanced data collection;

- High: empirical data were collected using a multi-method approach comprising multiple instruments.

The review showed that none of the evaluation designs were of a high quality, and only 37 per cent were of medium quality (meaning that empirical data was collected but the circumstances would allow for more advanced data collection). The remaining 63 per cent of identified evaluations were of low quality. These findings confirm a widely shared view that P/CVE evaluation is a relatively nascent area, where rigorous studies are hard to come by (e.g. Romaniuk 2015; Dawson et al. 2014; Veldhuis and Kessels, 2013).

1.2. Challenges to P/CVE evaluation

Existing literature includes extensive discussion of challenges to P/CVE evaluation (e.g. Romaniuk 2015; Chowdhury Fink et al. 2013; Romaniuk and Chowdhury Fink, 2012; Ris and Ernstorfer, 2017; Veldhuis and Kessels, 2013). Among the commonly cited difficulties are:

- **Definition of the concepts involved:** there is a lack of clarity on what radicalisation and violent extremism are, and what constitutes P/CVE. This is often further complicated by the lack of analysis on the specific problems that a given intervention is trying to address and the absence of a definition of what success looks like.
- **Devising indicators or metrics to measure relevant outcomes:** even if the success of an intervention has been defined, there are no established indicators or metrics to measure the outcomes of interest. Metrics, if used at all, vary considerably across interventions and validated scales or proxy indicators are not widely used.
- **Hard-to-reach, scarce and diverse populations of interest:** P/CVE

interventions aim to address a wide range of individuals and groups, ranging from those at risk of radicalisation through to those whose extremist views have already manifested in violence and those who need help to renounce violence. Some individuals may be disempowered or vulnerable to exploitation, and could be easily influenced by an evaluator. Others might be unscrupulous, manipulative or deceitful. The reluctance to engage these populations in an evaluation is a potential challenge and it requires trust to be built between those involved in the evaluation. These populations are often difficult to identify and small in size. As such, establishing a control or comparison group to compare the effects of an intervention with a counterfactual scenario might be problematic or infeasible.

- **Rarity of events:** despite capturing attention, acts of violent extremism are relatively infrequent. As such, it may be particularly difficult to demonstrate the effectiveness of interventions that aim to prevent such relatively rare events.
- **Dynamic contexts and politically driven agendas:** the situation and context in which P/CVE interventions occur changes continuously, limiting opportunities to establish a feedback loop through evaluation. Under uncertain and complex conditions, people tend to make sub-optimal decisions because of biases or cognitive shortcuts. When faced with potential threats, policymakers may disregard evaluation evidence and focus on one specific policy option, while failing to consider its possible unintended consequences. Moreover, violent extremism is ill-defined, multi-dimensional, and reliant on political and expert judgments rather than established evidence. Attempted solutions may have unintended – even

negative – outcomes, and as such P/CVE requires a multi-agency approach to develop solutions that can be continuously redefined and improved based on evaluation evidence and feedback.

Other challenges often discussed but not specific to P/CVE evaluations include a time lag before outcomes or impacts start to emerge, establishing causality and isolating other factors (attribution), and the sensitive nature of the subject.

Reflecting on these challenges evident in the CVE literature raises the question of how different evaluations in the CVE space really are from those addressing other complex interventions. Challenges that may appear unique to P/CVE may be replicated in other fields. For example, the complexity of the environment and the dynamic nature of interventions are also typical of stabilisation interventions (van Stolk et al. 2011; Stolk and Fazekas, 2013), while small numbers of cases and difficulties in finding a comparison are common in some types of crime (Harding et al. 2002) and in preventive public health programmes (Rugg et al. 1999; Tambuyzer 2010). Ill-defined concepts have proliferated in many policy areas, such as regional development (Markusen 1999). Finally, any evaluator is likely to be familiar with the pains of defining relevant outcomes and metrics.

So are P/CVE evaluations exceptional after all? Perhaps not. Perhaps what makes P/CVE evaluation unique is the combination of these challenges in a single area, but also the gravity associated with ‘getting it wrong’. However, some of the challenges cited above are relatively common, and as such many general evaluation approaches and methods can be successfully applied in P/CVE evaluations, as we explain below.

1.3. Previous efforts

Over the last 5 to 10 years, the volume of P/CVE evaluation literature has increased substantially. Below, we briefly discuss selected publications grouped under four broad categories: (i) analyses and assessments of P/CVE work; (ii) studies exploring different approaches to P/CVE evaluations; (iii) handbooks, guidance and manuals for P/CVE evaluations; and (iv) online P/CVE evaluation toolkits. While this selection is not exhaustive, it provides a broad overview of efforts in this area. What is striking when reading through these studies is that, as perhaps might be expected, a wide range of methods have been applied, or could be applied for evaluating P/CVE interventions.

1.3.1. Analyses and assessments of P/CVE work

In one of the earliest attempts to take stock of P/CVE evaluations, Romaniuk and Chowdhury Fink (2012) examined the development of the evaluation of counterterrorism policies, noting that policy evaluation was a new but expanding field. The authors recommended the use of mapping exercises, a database, training platforms for information sharing, and the development of online evaluation tools and a toolkit for policymakers and evaluators.

A comprehensive analysis of up-to-date factors in radicalisation, P/CVE interventions, evaluations and relations between these elements was provided by van Hemert et al. (2014) as part of the IMPACT Europe project. The authors classified and assessed evaluation methods and approaches used in the field and presented their findings in an interactive relational database that can be accessed via the IMPACT Europe evaluation toolkit to compare and eventually select the most appropriate evaluation methods for P/CVE interventions.

Another review of evaluation research on CVE, this time smaller in scope, was carried out by Romaniuk (2015). The author highlights a number of challenges for P/CVE evaluations and makes the case for systematising the understanding and practice of P/CVE, committing to its evaluation, and moderating expectations about its impacts. In a similar spirit, another study recommended the establishment of an institutionalised system to regularly monitor and evaluate the policies and measures in place (European Parliament, 2017).

1.3.2. Studies exploring different approaches to P/CVE evaluations

In 2010, Nelen et al. conducted a literature review on the ways of conducting empirical research on the impact of counterterrorism measures. The authors indicated that the evaluation of counterterrorism policy should be customised as a 'one-size-fits-all' approach to P/CVE evaluation was unrealistic. They also emphasised that any method that generates situation-specific knowledge ought to be embraced, but must meet minimum methodological requirements.

The Institute for Strategic Dialogue (2010) reviewed responses to violent radicalisation and made a number of suggestions on how evaluations should be organised and conducted. These included a multi-actor approach, a focus on process and outcomes and on learning and improvement, use of indicators, and embedding evaluations in a theory of change. Horgan and Braddock (2010) argued that a specific assessment technique, multi-attribute utility technology (MAUT, also known as multi-attribute evaluation) may be helpful in future empirical assessment of P/CVE initiatives. According to the authors, MAUT facilitated identification and weighting of the objectives held by a set of stakeholders, and the assessment of how well a given programme meets those goals

or objectives. Williams and Kleinman (2014) examined a utilisation-focused evaluation approach to address the question of how to assess the reduction of the risk of terrorism and attribute this reduction to a given programme. They provide a 'roadmap' for conducting such an analysis, and highlighted methodologically and theoretically challenges faced by evaluators.

Feddes and Galucci (2015) explored P/CVE evaluations between 1990 and 2014. The authors recommended empirical studies, using quantitative data where possible and a multi-method approach for evaluating programmes in challenging contexts. The importance of context was also emphasised by Gielen (2017), who advocated realist evaluation to assess counterterrorism programmes.

Another publication by Ris and Ersntofer (2017) explored the challenges of designing monitoring and evaluation methods for P/CVE programmes and advocated methods used in conflict prevention that could be applicable to P/CVE programmes. Similarly, Davies et al. (2017) explored the transferability and applicability of gang-related evaluation methodologies to P/CVE.

Again, considering the approaches, methods and techniques found in the literature, only some of these (e.g. MAUT) suggest that P/CVE evaluations are exceptional, while others draw on more traditional approaches (theory of change, realist evaluation, mixed methods, etc.). This is not surprising given that many well-known evaluation approaches and methods can also be used in P/CVE evaluations.

1.3.3. Handbook, guidance and manuals for P/CVE evaluations

The applicability of established evaluation approaches is well illustrated by a growing number of handbooks, guides and manuals

on P/CVE evaluation, most of which draw on traditional rather than P/CVE-specific methods. These resources are briefly outlined below.

- A handbook for practitioners titled 'Learning and Adapting' highlights the value of monitoring and evaluating P/CVE programmes. The handbook introduces lessons learned from P/CVE evaluation in other countries and emphasises that the continued strength of P/CVE approaches depends on the ability to demonstrate that these projects and interventions deliver impact, insights and return on investment (Dawson et al. 2014).
- A US Department of State (2016) guide on 'Monitoring for CVE' helps document progress made in fighting violent extremism. The guide focuses on how to facilitate contextual P/CVE analysis, how to use monitoring as a management tool and how to prepare for an evaluation.
- The handbook of the Institute for Strategic Dialogue aims to support proactive response to extremist propaganda with counter-narrative campaigns, and is intended as a beginner's guide for those with little or no previous experience of counter-narrative campaigning (Tuck and Silverman, 2016). The handbook covers the main stages of creating, launching and evaluating an effective counter-narrative campaign.
- In conjunction with DHS Community of Office Partnerships, RAND Corporation developed an evidence-based model to help community P/CVE programmes conduct self-evaluations. The RAND Program Evaluation Toolkit for Countering Violent Extremism helps overcome common challenges to evaluating P/CVE

programmes by developing a logic model to show connections between resources, activities, outcomes, evaluation measures, and the need the programme addresses in its community. It then helps users design an evaluation that is appropriate for their intervention type and the available resources and expertise (Todd et al. 2017).

Finally, the IMPACT Europe evaluation toolkit⁴ has been developed as an interactive step-by-step online resource to promote and support P/CVE evaluations. The final report presents the results, outputs and lessons learned from the IMPACT Europe project (Marret et al. 2017). The toolkit is a knowledge-management tool designed to support P/CVE practitioners in designing and implementing better interventions and evaluations. It is tailored to the overall P/CVE context and to concrete needs on the ground. The toolkit adds value to existing practice in several ways. Notably, it provides a single entry point to comprehensive information about radicalisation factors and P/CVE interventions and evaluations.

1.4. About this collection

This document derives from the work carried out under the IMPACT Europe project. It synthesises the collection of methods and approaches included in the IMPACT Europe evaluation toolkit.

It provides an overview of data collection methods and evaluation designs. It helps inform decisions about the method or design that will be suitable to answer specific information needs, and which will be appropriate given the context and practical considerations of an evaluation. While the list of methods is not exhaustive, the ones presented here were selected based on

the breadth of data collection and analysis covered, as well as their appropriateness for CVE programmes.

This collection has been developed with the P/CVE context in mind and with the aim of tailoring descriptions of methods to specific characteristics of work in this field. Where examples of the methods featured this collection being used in P/CVE evaluations were identified, these have been included to illustrate when and how these methods can be effectively applied.

This document also provides a comprehensive discussion on ethical considerations in P/CVE evaluation more broadly, as well as in the application of specific evaluation methods. Offering a wealth of resources, examples and further reading, this document is a valuable guide for those who want to enhance their P/CVE evaluation toolbox.

List of references

Chowdhury Fink, N., Romaniuk, P. and Barakat, R. 2013. *Evaluating Countering Violent Extremism Programming: Practice and progress*. Center on Global Counterterrorism Cooperation.

Davies, M., Warnes, R. and Hofman, J. 2017. *Exploring the transferability and applicability of gang evaluation methodologies to counter violent radicalisation*. Santa Monica, Calif: RAND Corporation. As of 18 February 2018: https://www.rand.org/pubs/research_reports/RR2120.html

Dawson, L., Edwards, C. and Jeffray, C. 2014. *Learning and Adapting: The Use of Monitoring and Evaluation in Countering Violent Extremism – A Handbook for Practitioners*. Royal United Services Institute.

European Parliament. 2017. *The European Union's Policies on counterterrorism, relevance,*

coherence and effectiveness. European Parliament.

Feddes, A.R. and Galucci, M. 2015. 'A Literature Review on Methodology used in evaluating effects of Preventive and De-radicalisation interventions'. *Journal for Deradicalization* (5), pp.1-27.

Gielen, A.J. 2017. 'Countering Violent Extremism: A Realist Review for Assessing What Works, For Whom, in what circumstances and how?'. *Terrorism and Political Violence*, May 2017:1–19.

Harding, D. J., Fox, C., & Mehta, J. D. 2002. 'Studying rare events through qualitative case studies: Lessons from a study of rampage school shootings.' *Sociological Methods and Research*, 31(2), 174-217.

Horgan, J. and Braddock, K. 2010. 'Rehabilitating the Terrorists?: Challenges in Assessing the Effectiveness of De-radicalization Programs'. *Terrorism and Political Violence*, 22:267-291. As of 18 February 2018: www.start.umd.edu/start/publications/Derad.pdf

Institute for Strategic Dialogue. 2010. *Comparative Evaluation Framework for Counter Radicalisation*. PPN Working Paper. As of 18 February 2018: <https://www.counterextremism.org/resources/details/id/9/ppn-working-paper-comparative-evaluation-framework-for-counter-radicalisation>

Markusen, A. 1999. 'Fuzzy Concepts, Scanty Evidence, Policy Distance: The Case for Rigour and Policy Relevance in Critical Regional Studies', *Regional Studies*, 33:9, 869-884.

Marret, J.L., Bellasio, J. van den Berg, H., van Gorp, A., van Hemert, D., Leone, L., Meijer, R., Warnes, R., Van Wonderen, R. 2017. *Final report providing background for using and further developing the validated toolkit, Innovative Methods and Procedures to Assess Counter-violent-radicalisation Techniques in Europe*.

IMPACT Europe project. As of 18 February 2018: <http://www.impacteurope.eu/>

Nelen, H., Leeuw, F. and Bogaerts, S. 2010. *Antiterrorismebeleid en evaluatieonderzoek*. Boom Juridische Uitgevers.

Ris, L. and Ernstofer, A. 2017. *Borrowing a Wheel: Applying Existing Design, Monitoring and Evaluation Strategies to Emerging Programming Approaches to Prevent and Counter Violent Extremism*. Peacebuilding Evaluation Consortium.

Romaniuk, P. and Chowdhury Fink, N. 2012. *From input to impact: Evaluating Terrorism Prevention Programs*. Center on Global Counterterrorism Cooperation.

Romaniuk, P. 2015. *Does CVE Work? Lessons Learned from the Global Effort to Counter Violent Extremism*. Indiana: Global Center on Cooperative Security.

Rugg, D., Buehler, J., Renaud, M., Gilliam, A., Heitgerd, J., Westover, B., Wright-Deaguero, L., Bartholow, K. and Swanson, S., 1999. 'Evaluating HIV prevention: A framework for national, state and local levels'. *American Journal of Evaluation*, 20(1), pp.35-56.

Tambuyzer, E. 2010. 'Rare diseases, orphan drugs and their regulation: questions and misconceptions'. *Nature Reviews Drug Discovery*, 9(12), 921-929.

Todd, H., Matthews, M., Ramchand, R., Beaghley, S., Stebbins, D., Kadlec, A., Brown, M.A., Kofner, A. and Acosta, J.D. 2017. *RAND Program Evaluation Toolkit for Countering Violent Extremism*. Santa Monica, CA: RAND Corporation. As of 18 February 2018: <https://www.rand.org/pubs/tools/TL243.html>

Tuck, H., and Silverman, T. 2016. *The Counter-Narrative Monitoring and Evaluation Handbook*. Institute for Strategic Dialogue.

US Department of State. 2016. *Monitoring for CVE: Documenting progress on the fight against violent extremism*. US Department of State.

van Hemert, D., van den Berg, H., van Vliet, T., Roelofs, M. and in 't Veld, M.. 2014. *Synthesis report on the state-of-the-art in evaluating the effectiveness of counter-violent extremism initiatives*. IMPACT Europe project.

van Stolk, C., Ling, T., Reding, A., Bassford, M. 2011. *Monitoring and evaluation in stabilisation interventions: Reviewing the state of the art and suggesting ways forward*. Santa Monica, CA: RAND Corporation. As of 18 February 2018: https://www.rand.org/pubs/technical_reports/TR962.html

van Stolk, C. and Fazekas, M. 2013. *How Evaluation Is Accommodated in Emergency Policy Making: A Comparison of Post-9/11 Emergency Legislation in the United Kingdom and the United States*. Published in: *Evaluation and Turbulent Times: Reflections on a Discipline in Disarray* (Comparative Policy Evaluation, Vol. 20) / Jan-Eric Furubo, Ray C. Rist, Sandra Speer, eds. (New Brunswick, NJ: Transaction Publishers, 2013), Chapter 9, p. 161-177.

Veldhuis, T.M. and Kessels, E.J.A.M. 2013. *Thinking before Leaping: The Need for More and Structural Data Analysis in Detention and Rehabilitation of Extremist Offenders*. ICCT Research Paper.

Williams, M.J. and Kleinman, S.M. 2014. 'A Utilization-Focused Guide for Conducting Terrorism Risk Reduction Program Evaluations'. *Behavioral Sciences of Terrorism and Political Aggression*, 6(2).

CHAPTER TWO

Developing this collection and user guide

Joanna Hofman and Megan Sim

This chapter presents the process of identifying and selecting relevant approaches and methods included in this collection.

2.1. Selection of methods

To gain useful lessons from evaluations, policymakers and practitioners need to understand not only whether or not an intervention had any effect but also **why** and **how** effects were realised. A range of evaluation approaches and methods, carefully tailored to each intervention, might be needed to address these questions. With this in mind, methods were included in this collection on the grounds of the breadth of data collection and analysis approaches covered, as well as their appropriateness for P/CVE programmes. The evaluation methods were selected in three stages, which are outlined below.

First, to select relevant evaluation approaches and methods we identified those that have already been used in evaluations of P/CVE

interventions. A description of the procedure to identify P/CVE evaluations and results of the review (including information about methods and designs used in each P/CVE evaluation) was carried out as part of the IMPACT Europe project and was detailed by van Hemert et al. (2014). The most common approaches and methods included: cross-sectional studies, observation techniques, interviews, focus groups, surveys, data mining, and combinations of these. While not widespread, a small number of longitudinal studies, quasi-experimental approaches and case-studies were also identified.

Second, we reviewed evaluation methods that have been used in evaluations of gang desistance interventions, a field which shares some characteristics with P/CVE work, and where evaluation of interventions is more mature. The method of identifying gang evaluations, including the assessment of applicability and transferability of methods to the P/CVE field, is presented in Davies et

al. (2017). This study pointed to longitudinal studies as more suitable to capturing impact and encouraged more experimental and quasi-experimental designs, such as randomised control trials, comparison groups, etc. However, these could be more difficult to translate to P/CVE contexts.

Third, in order to select evaluation methods we relied on in-house evaluation knowledge, academic sources, and other guidelines and documents, including those published by:

- European institutions (European Commission, 2015) and international organisations, such as the United Nations (2013) and World Bank (IEG-World Bank, 2007; Gertler, 2016);
- National governments, including the UK Treasury (HM Treasury, 2011), the UK Department for International Development (2005; Stern et al. 2012); and
- Research institutes (Ling and Villalba van Dijk, 2009).

As a result of this process, we arrived at an initial list of 43 methods. To provide a comprehensive yet manageable selection we discussed the list within the IMPACT Europe team and shortlisted 24 methods through consensus. Our choice was partly driven by pragmatic considerations, including a desire for consistency with the findings of IMPACT Europe, and resource limitations that did not allow us to expand the list to include all the methods and approaches identified as suitable. We also note that this selection should not be considered as ultimate or exhaustive (see section 2.4. on limitations).

The final selection of the 24 most common and applicable methods reflects the wide range of instruments that can be used for data collection and analysis to make evaluations more robust. Our selection was guided by the principle of encouraging the use of instruments

which, where possible, go beyond simply describing what happened by assessing whether the intervention has had effects, and the magnitude of these effects. We do not suggest that the aim of understanding cause and effect commits us to the exclusive use of experimental and statistical designs. While they are certainly useful in some circumstances, a wider palette of methodological options is needed. Often multiple causal mechanisms coexist and we may need mixed-methods or theory-based evaluations to disentangle these.

However, even with the best instruments available to evaluate them, since many P/CVE programmes share the same aims as other programmes and actions (a more cohesive society, community engagement, reduced criminality, etc.), there may be limits to how far we can attribute changed outcomes to the effects of the programme. In this situation the goal is to determine the strength of the evidence that the programme at least made a contribution to the intended outcome.

Furthermore, as part of their inherent complexity, P/CVE programmes may adapt to circumstances or new information, meaning that the programme design at the end of implementation is different to the original design. A further complication is that effects may only be measurable after the life of the programme or may be non-linear (for example, there may be no measurable change until a tipping point is reached, after which there is dramatic change).

These challenges are outlined here not to suggest that P/CVE interventions are exceptional or cannot be well-evaluated, but rather to emphasise the need for a range of approaches, carefully tailored to the needs of each intervention. In developing our instruments, we did not set out to produce a simple 'cook book' for evaluation design. Rather, this document is intended to provide a basis for rational and informed choices.

2.2. Method and research design summaries

We have developed a brief and tailored description of each method/research design, drawing on widely available source material, with practical information and specific consideration of how each can be implemented in a P/CVE context. In order to allow for comparisons between the methods, we followed a standardised template which structured information as follows:

- Brief description
- Purpose
- When (not) to use it
- Step-by-step application
- Ethical considerations
- Further resources.

Additionally, we searched for examples of P/CVE evaluations which used any of the selected methods and how it was done. For instance, in order to identify examples of how individual methods have been applied in P/CVE, we reviewed evaluations identified by van Hemert et al. (2014) and the Radicalisation Awareness Network's collection of approaches, lessons learned and practices (RAN, 2018). These examples, where identified, were included in the descriptions of methods. Not all methods feature these examples, which reflects a number of considerations: that not all methods have been commonly used in P/

CVE evaluations, that P/CVE evaluations do not consistently outline the methods and approaches used, and that some evaluations that used methods presented here were not identified or not publicly available. We also acknowledge that the following chapters connect methodologies and approaches to the world of P/CVE in somewhat varied ways. In some cases the link is well explained and supported by examples, while in others (usually those less commonly used or more difficult to apply in P/CVE contexts) this connection might be more tangential.

2.3. Classification of methods

The methods presented in this collection vary significantly in scope and nature. They include techniques for collecting data (such as interviews) and for analysing and presenting data (e.g. descriptive statistics), and approaches to making design choices and framing an evaluation that include data collection and analysis (e.g. realist evaluation). As such, the methods and approaches presented here are not equivalent to each other.

There is a wide variety of approaches to classifying evaluation methods (e.g. Ling and Villalba van Dijk, 2009). We did not cluster methods based on established theoretical approaches. Instead, we provide practical information about each method which can be classified based on a number of characteristics, as outline in Table 1.

Table 1. Classification of methods included in this collection

Characteristic	Categories	Description
Type	Design or method	Evaluation design refers to the plan for conducting the evaluation. Evaluation method refers to the way in which data is collected and/or analysed.
Approach	Qualitative or quantitative	Qualitative approaches broadly include historical, intuitive and observational approaches. Quantitative approaches use statistical, mathematical or computational techniques.
Focus	Economic/ impact or mechanism/ process	An economic evaluation of an intervention refers to an analysis of its financial costs, while an impact evaluation examines the results (and thus impact) of the intervention.
		A focus on the intervention mechanism refers to an examination of why it works, while a process evaluation examines whether the intervention components were implemented and how successful the implementation was.
Data used	Primary or secondary	Primary data refers to data that has been collected directly by the evaluator. Secondary data refers to data that has been collected by someone other than the evaluator.
Purpose	Data collection or data analysis	Data collection refers to the process of systematically gathering and measuring information of interest for the evaluation. Data analysis refers to the process of systematically describing, summarising and evaluating information collected as part of an evaluation.

The collection of evaluation methods contains descriptions of 24 methods. Each of the 24 methods presented has been tailored to the P/CVE context to give readers more specific information on how each method can be applied in practice. An overview of the research methods and their possible uses is available in Table 2. Rather than grouping these methods according to certain uses or approaches, we present them in an alphabetical order.

It is important to note that evaluation often relies on several methods described in this collection, and may also depend on methods not described here. The central question about designing an evaluation is whether the methods/design allow one to answer research questions. Decisions around that also depend on other factors, including:

purpose of evaluation, evaluation questions, information needs and the level of ambition for the evidence to be collated, context in which the evaluation is carried out, and pragmatic considerations (including availability of secondary data, access to stakeholders, cost and time implications, etc.).

2.4. Limitations

There are a number of limitations to this collection of evaluation methods.

First, evaluation is a dynamic practice where new methods and approaches are being developed and tested. As such, other and innovative methods might also be suitable and worth considering when planning a P/CVE evaluation.

Table 2. Overview of evaluation approaches and methods and their uses

	Is it a		Is it		Can it be used for		Does it involve		Is it used for	
	Design	Method	Qualitative	Quantitative	Impact / economic evaluation	Process / mechanism evaluation	Primary data	Secondary data	Collection of data	Analysis of data
1. Case studies	Y	Y	Y	Y	Y	Y	Y	Y		Y
2. Comparisons / benchmarking	Y	Y	Y	Y	Y	Y	Y	Y		Y
3. Contribution analysis	Y	Y	Y		Y		Y	Y		Y
4. Cost-benefit analysis		Y		Y	Y		Y	Y		Y
5. Cost-effectiveness analysis		Y		Y	Y		Y	Y		Y
6. Cross-sectional data analysis	Y	Y		Y	Y	Y	Y	Y		Y
7. Data mining		Y		Y	Y			Y		Y
8. Descriptive statistics		Y		Y	Y	Y	Y	Y		Y
9. Desk-based research / literature review		Y	Y		Y	Y		Y	Y	Y
10. Focus groups		Y	Y		Y	Y	Y		Y	
11. Interviews		Y	Y		Y	Y	Y		Y	
12. Logic models / theory of change	Y	Y	Y		Y	Y	Y	Y	Y	Y
13. Longitudinal	Y		Y	Y	Y	Y	Y	Y	Y	
14. Meta analysis	Y	Y		Y	Y			Y		Y
15. Network analysis		Y	Y	Y	Y	Y	Y	Y		Y
16. Objectives and options analysis		Y	Y		Y	Y	Y	Y	Y	Y
17. Observation techniques / ethnography	Y	Y	Y		Y	Y	Y		Y	Y
18. Policy scientific approach	Y	Y	Y		Y	Y	Y	Y	Y	Y
19. Qualitative data analysis		Y	Y		Y	Y	Y	Y		Y
20. Quasi-experimental designs	Y			Y	Y		Y	Y		Y
21. Randomised control trials	Y			Y	Y		Y		Y	Y
22. Realist evaluation	Y	Y	Y		Y	Y	Y	Y	Y	Y
23. Stakeholder analysis		Y	Y		Y	Y	Y	Y		Y
24. Surveys		Y	Y	Y	Y	Y	Y		Y	

Second, the list of methods is not exhaustive. It aims to showcase 24 relevant and commonly applied approaches and methods which, as explained earlier, vary significantly in scope and nature. We acknowledge that there are many more methods to draw from.

Third, this resource is intended to give readers an overview of the methods and how they can be used, but it does not replace practical experience. Thus, a practitioner who is unfamiliar with a particular method or design may have to use the 'further reading' suggestions presented in the respective sections to learn more about how to employ the method in practice. Description of the methods was intentionally kept limited so as not to overwhelm readers with information. Instead, it provides readers with general information about the method to help them understand which method would be suitable for their own evaluation.

Thirdly, as emphasised earlier, often a range of approaches and multiple methods, carefully tailored to the needs of each intervention, is required to deliver a robust evaluation of a P/CVE intervention. Relying on a single method in any evaluation bears a risk of providing skewed or partial answers to evaluation questions. This collection serves as a menu for making informed choices when developing a bespoke evaluation design, and we encourage users to consider mixed-method approaches.

Fourthly, as noted earlier, the choice of methods should be guided by the purpose of an evaluation and evaluation questions one aims to address. Developing an appropriate framework for the collection and analysis of data requires careful consideration. We hope that this collection and the methods included will help users to make these choices.

List of references

- Davies, M., J. Hofman and R. Warnes. 2017. *Exploring the transferability and applicability of gang evaluation methodologies to counter violent radicalization*. Santa Monica, CA: RAND Corporation. As of 18 February 2018: https://www.rand.org/pubs/research_reports/RR2120.html
- Department for International Development. 2005. *Guidance on Evaluation and Review for DFID Staff*. Evaluation Department, London: Department for International Development.
- European Commission. 2015. *Better Regulation Toolbox accompanying the Better Regulation Guidelines*. SWD(2015) 111 final, Commission Staff Working Document. Luxembourg: The Publications Office of the European Union.
- Gertler, P.J., Martinez, S., Premand, P., Rawlings, L.B. and Vermeersch, C.M., 2016. *Impact evaluation in practice*. Washington, D.C.: World Bank.
- Independent Evaluation Group–World Bank. 2007. *Sourcebook for Evaluating Global and Regional Partnership Programs. Indicative Principles and Standards*. Washington, D.C.: IEG-World Bank.
- Ling, T. and L. Villalba van Dijk. 2009. *Performance Audit Handbook: Routes to effective evaluation*. Santa Monica, CA: RAND Corporation. As of 18 February 2018: https://www.rand.org/pubs/technical_reports/TR788.html
- Radicalisation Awareness Network (RAN). 2018. *Preventing Radicalisation to Terrorism and Violent Extremism Approaches and Practices*. RAN Collection of Approaches and Practices. Luxembourg: The Publications Office of the European Union.

Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., Befani, B. 2012. *DFID Working Paper 38. Broadening the range of designs and methods for impact evaluations*. DFID, London.

United Nations Evaluation Group. 2013. *Impact Evaluation in UN Agency Evaluation Systems: Guidance on Selection, Planning and Management*. UNEG Impact Evaluation Task Force.

van Hemert, D., van den Berg, H., van Vliet, T., Roelofs, M. and in 't Veld, M.. 2014. *Synthesis report on the state-of-the-art in evaluating the effectiveness of counter-violent extremism initiatives*. IMPACT Europe project.

HM Treasury. 2011. *The Magenta Book: Guidance for evaluation*. London: HM Treasury.

CHAPTER THREE

Ethical considerations

Anke van Gorp

Below we discuss ethical issues when conducting evaluations of P/CVE interventions. The European Convention on Human Rights (especially articles 8, 9, 10, 11 and 14) and the Charter of Fundamental Rights of the European Union (especially articles 6, 7, 8, 10, 11, 12, 20 and 21) should be a baseline for all P/CVE interventions designed and conducted in the European Union. This means that both the intervention and the evaluation need to meet these standards. We note that this is not unique to P/CVE – the requirements set out below are a practical guideline to complying with these human rights during any evaluation. For example, informed consent about participating in an evaluation and confidentiality of information can be seen as way to protect the privacy of respondents and a way to meet data protection requirements regardless of the policy area in which an evaluation takes place. Below we will focus on the ethical issues connected to **evaluation** of P/CVE interventions and not on the P/CVE interventions themselves. However, many of these issues are considered good practice more generally, rather than specifically in the

P/CVE field. Given that there are few generic evaluation resources that cover ethical issues systematically, we hope this general guidance will also be helpful for P/CVE practitioners.

3.1. General ethical issues surrounding evaluation

In the evaluation of any policy or intervention, there are always interests at stake, which are sometimes competing: those people and organisations *performing the programme* under evaluation; people/organisations *funding* the programme; *programme participants*; and *evaluators*. Although all these stakeholders have roles in the evaluation, some have more power than others. This means that choices of evaluation methods may be contested, if they favour some groups over others.

In designing an evaluation, one of the first questions is: who will conduct the evaluation? This may be an internal party, an external party or a part of an organisation that is not directly involved in the programme under evaluation but part of the organisation that is performing the intervention. An internal evaluator knows

a programme and the people working on it. This might be beneficial for gaining access but it might also create bias in the evaluator. Hiring an external evaluator might seem more objective, but the fact that an evaluator is paid by the organisation whose programme is being evaluated means that the external evaluator may also be perceived as being biased (Ariely, 2012), yet that belies the reality of independent evaluation, which is frequently undertaken on a contractual basis.

Another important question is what type of evaluation is chosen. In the list of evaluation methods and approaches provided in this paper, different methodological and ethical issues are raised. In choosing a method, a choice is made about what type of evidence is gathered and what type of evidence is not pursued, which then has later impacts on how an intervention is framed and presented.

Questions about what constitutes 'proof' of whether an intervention is effective are highly contested and answers to these questions vary to some extent between scientific fields. Although these questions are of a methodological nature, they also contain an ethical dimension because they are related to what type of data is gathered and who can (and who cannot) provide the necessary data. In this paper, these questions are raised in the section on ethical considerations included in each chapter.

3.2. Ethical issues when conducting an evaluation

In this section we outline the types of ethical issues faced by evaluators during the course of an evaluation. Many of these issues, for instance the importance of treating respondents and their data with respect, are similar to those arising in research more generally. Within research ethics, norms have

been developed about how to treat participants in social science research (see for example Dench et al. 2004; Oliver, 2010). As Guillemin and Gillam (2004) have argued, ethics in social science research cannot be completely captured in rules and norms. Besides rules about, for example, informed consent, social scientists need to be ethically sensitive. Although not every ethical issue can be prevented or captured in rules and norms, the following norms are considered good practice in social science:

- **Informed consent:** whenever possible, respondents in interviews, focus groups, surveys and observations 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. For requirements on informed consent and also indirect ways to ensure voluntary participation in social science research. With regard to informed consent, researchers should provide information in a language and a format that is understandable for respondents. This means that information for experts is different in language and lay-out from information for adolescents or children. Researchers need to provide the following information to respondents as part of the process of obtaining their explicit consent:
 - The goal of the research;
 - Clarification that participation in the evaluation is voluntary, and wording such as: 'If you choose not to participate, you will not be penalised, nor will you forgo any benefits to which you are otherwise entitled';

- Explanation that respondents can quit at any time during the interview, workshop or survey;
 - How respondents' data are used and for what (for example types of publications);
 - Who gets access to respondents' data;
 - To whom respondents can turn if they have questions or want (some part of) their personal data removed from the study (this means that their personal data is withdrawn from that point in time – there is no need or possibility to retract personal data from analyses that are already published);
 - Details of approval of the evaluation by an ethical committee (if such an approval has been obtained) and, if the organisation has a Data Protection Officer, his or her contact information.
- **Inclusion of minors:** the inclusion of minors in social research varied 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. This will depend on the type of respondents and the method used.
 - **Minimise harm and maximise good:** evaluations should be designed, conducted and disseminated with the aim of benefiting public good. Careful thought should be given to the potential outcomes of evaluation and how findings might be used. Individual evaluation participants (and the wider social groups or organisations to which they belong) and evaluators themselves ought to have their physical, social and psychological wellbeing protected. As such, evaluators need to anticipate and guard against any possible harmful consequences (to participants and evaluators) of engaging in an evaluation. If there is a high potential risk of distress, evaluators should explore whether there are other means to obtain the information that is needed – only if there are no other ways and the information is clearly needed data could be collected. Where risks to participants and/or researchers are unavoidable as part of the research, robust risk assessment and management procedures should be in place (including how the evaluators will respond to signs of distress and the provision of post-participation support).
 - **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. It is difficult to judge what questions could emotionally burden respondents as, for example, questions about family life do not have significant impact on some respondents whereas others (who have experienced difficulties in family life) may find such questions emotionally burdensome. During interaction with respondents, evaluation researchers should be sensitive

to the emotional burden for respondents (see for example Guillemin and Gillam, 2004). It is impossible to completely predict the emotional burden of research, and interviewers should be sensitive to this. With regard to subjects such as violence, protocols can be designed to help researchers know what to do if respondents tell them they are a victim of any type of violent act.

- **Gathering information about individuals other than respondents:** whether or not individuals will become respondents in an evaluation, information about them is gathered if other individuals cooperating with them are respondents. In evaluation research, it can be necessary to not only ask what a respondent him/herself thinks or has done, but also what the respondent believes others think or have done (Borgatti and Molina, 2003). Borgatti and Molina (2003) claim that this cannot be changed and is acceptable, though the interests of individuals who are not respondents but still appear as an actor in the evaluation should be taken into account when writing up.
- **Use of a control group:** if experiments or quasi-experiments are used to evaluate effectiveness of a programme, there are issues with regard to the use of control groups. There are questions as to whether the control group is getting no support from the programme, delayed support (so-called 'waiting list' control group), or whether 'business as usual' is deemed unacceptable. If people are already radicalised and have used violence then a strategy of waiting without any support might be deemed unacceptable – see for example Street and Luoma (2002) and Saks et al. (2002).

With regard to the ethical issues in evaluation, a paragraph on ethics is included in the description of every method (see Chapters 4–27).

3.3. Ethics in the context of interventions in the field of de-radicalisation, counter-radicalisation and counter-polarisation

All personal information about participants in programmes is very sensitive, but this may be particularly the case for interventions attempting to counter violent extremism. This means that the identity of participants must remain confidential and that care should also be taken that participants cannot be indirectly identified. Moreover, in some evaluations, evaluators might want to ask for sensitive information from participants in P/CVE interventions, for example about their political or religious views. Under the GDPR, this information falls under special categories of personal data which can only be gathered, at least by evaluators who are outside of government, with explicit consent and should be securely stored. In-depth interviews could lead to the disclosure to the evaluator/interviewer of problems or dangerous family situations. An evaluator or interviewer needs to be prepared to react to such disclosures (Guillemin and Gilliam, 2004). The evaluators of certain P/CVE programmes, for example P/CVE programmes with individuals who have been convicted for extremist violence or terrorism in prison, might also experience risks. These risks may be physical but also emotional or professional (see for example Jipson and Litton, 2000).

3.4. Further reading on codes of ethics/conduct

When evaluating P/CVE interventions, one can make use of existing codes of conduct and we recommend two as further reading in this area. Below are guidelines from the United Nations Development Programme (UNDP) and the

American Evaluation Association. UNDP (2011) identifies a set of guiding principles, which hold that an evaluation should be:

- **Independent:** management must not impose restrictions on the scope, content, comments and recommendations of evaluation reports. Evaluators must be free of conflicts of interest.
- **Intentional:** the rationale for an evaluation and the decisions to be based on it should be clear from the outset.
- **Transparent:** meaningful consultation with stakeholders is essential for the credibility and utility of the evaluation.
- **Ethical:** evaluation should not reflect personal or sectoral interests. Evaluators must have professional integrity, respect the rights of institutions and individuals to provide information in confidence, and be sensitive to the beliefs and customs of local social and cultural environments.
- **Impartial:** removing bias and maximising objectivity are critical for the credibility of the evaluation and its contribution to knowledge.
- **Of high quality:** all evaluations should meet minimum quality standards (in UNDP's case, those defined by its Evaluation Office).
- **Timely:** evaluations must be designed and completed in a timely fashion so as to ensure the usefulness of the findings and recommendations.
- **Used:** evaluation is a management discipline that seeks to provide information to be used for evidence-based decision making. To enhance the usefulness of findings and recommendations, key stakeholders should be engaged in various ways in the conduct of the evaluation.

The American Evaluation Association (2004) identifies five guiding principles:

- **Systematic inquiry:** evaluators conduct systematic, data-based inquiries about whatever is being evaluated.
- **Competence:** evaluators provide competent performance to stakeholders.
- **Integrity/honesty:** evaluators ensure the honesty and integrity of the entire evaluation process.
- **Respect for people:** evaluators respect the security, dignity and self-worth of respondents, programme participants, clients, and other stakeholders with whom they interact.
- **Responsibilities for general and public welfare:** evaluators articulate and take into account the diversity of interests and values that may be related to the general and public welfare.

List of references

- American Evaluation Association (2004). As of 29 November 2014: <http://www.eval.org>
- Ariely D. 2012. *The honest truth about dishonesty*, New York: HarperCollins Publishers.
- Borgatti S.P. and Molina J.L. 2003. Ethical and Strategic Issues in Organizational Social network Analysis, *The Journal of Applied Behavioral Science*, 39 (3), 337-349.
- Council of Europe. 1950. *European Convention for the Protection of Human Rights and Fundamental Freedoms*, as amended by Protocols Nos. 11 and 14, 4 November 1950, ETS 5. As of 10 July 2017: <http://www.refworld.org/docid/3ae6b3b04.html>
- Dench, S., Iphofen, R. and Huws, U. 2004. *An EU code of Ethics for Socio-Economic Research*, IES report 412, The Institute for employment studies, Brighton.

European Union. 2012. Charter of Fundamental Rights of the European Union, 26 October 2012, 2012/C 326/02. As of 10 July 2017:

<http://www.refworld.org/docid/3ae6b3b70.html>

Guillemin M. and Gillam L. 2004. Ethics, Reflexivity, and “Ethically Important Moments” in Research, *Qualitative Inquiry*, 10(2), 261-280.

Jipson A.J. and Litton C.E. 2000. Body, Career and Community; the implications of researching dangerous groups in: Linkogle S. and Lee-Treweek G. *Danger in the field: Ethics and risks in social research*, Florence, KY, US; Routledge.

Oliver P. 2010. *Student's guide to research ethics, second edition*, Maidenhead: Open University Press.

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April

2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). OJ L 119, 4.5.2016, p. 1–88.

Saks E., Jeste D.V., Granholm E., Palmer B.W. and Schneiderman L. 2002. Ethical issues in psychosocial interventions research involving controls, *Ethics & behaviour*, 12(1), p 87-101.

Street L. and Luoma J. B. 2002. Control groups in psychosocial intervention research: ethical and methodological issues, *Ethics & behaviour*, 12(1), p 1-30.

UNDP. 2011. *The Evaluation Policy of UNDP*, New York: Executive Board of the United Nations Development Programme and of the United Nations Population Fund.



CHAPTER FOUR

Case studies

Kate Cox and Anke van Gorp

See also:



Example(s) of use

The International Centre for the Study of Radicalisation and Political Violence (ICSR) used case studies in its 2010 report 'Prisons and Terrorism: Radicalisation and De-radicalisation in 15 Countries'. Through case study analysis of 15 countries, the report assessed the role that prisons play in radicalising and de-radicalising people. The case studies highlighted practices to help policymakers identify new approaches and avoid costly mistakes (Neumann, 2010).

The 2008 'Radicalisation, Recruitment and the EU Counter-radicalisation Strategy' report also drew upon case study research. This report is part of the European Commission-funded 'Transnational Terrorism, Security and the Rule of Law' project, which ran from 2006 to 2009. The research used two case studies to examine the measures taken by the UK government and Amsterdam local authorities to counter-radicalisation (Staun, 2008).

4.1. Brief description

A case study is an empirical inquiry that investigates a subject or issue within its real-life context, assuming that the context plays an important explanatory role in understanding the subject matter. A unit of analysis can be a country, an organisation within the country, a programme run by the organisation or a client of the programme. The researcher typically selects a limited number of case studies

because they can be resource intensive or are intended as illustrative of a range of issues. For instance, a set of case studies may look at a few individual clients to better understand the effectiveness of a P/CVE programme that aimed to facilitate their departure from radicalised groups operating in different regions or locations.

4.2. Purpose

Case studies allow the researcher to collect detailed data on rare or complex phenomena which can form a strong evidence base for an evaluation. They are also an effective way of drawing comparisons between different actors within a single case, between cases, and between groups across cases. For instance, an evaluation using case studies to assess the effectiveness of a P/CVE intervention implemented in different regions or locations could help identify factors or circumstances that have contributed to some regions being more successful than others. Case studies can support understanding of how different elements (such as varying contexts, implementation arrangements and other factors) fit together and how they create the pathways by which outcomes or impacts occur.

4.3. When to use it

There are several different types of case study. A case study can be exploratory (create new knowledge), descriptive (outline the phenomenon and its context), constructive (solve a problem), instrumental (facilitate understanding of something else), explanatory (explain causal links in interventions), or confirmatory (test a hypothesis with empirical evidence) (US General Accounting Office, 1990).

Case studies are particularly useful when the research aims to explain a real-life phenomenon which is complex and may require various methods to collect and analyse data. In fact, case studies can be considered an analytical method and a research design.

A case study is especially helpful in evaluating programmes that are unique, when an established programme is implemented in a new setting, when a unique outcome warrants further investigation, or when a programme occurs in an unpredictable environment.

4.3.1. Why?

Advantages of case studies include:

- **Understanding of change processes:** case studies are useful for understanding the process and dynamics of change. By closely documenting and interpreting events, case studies can determine the critical factors in programme implementation and analyse the links between them (e.g. actors involved in a P/CVE programme implementation and their roles, and how these roles changed over time).
- **Detailed information:** case studies provide detailed information concerning a specific policy, programme, system or individual (e.g. de-radicalisation policy targeting youth, training policy for teachers to identify individuals at risk of radicalisation, characteristics of radicalised group leaders, etc.) through an in-depth data collection process.
- **Multiple methods and perspectives:** using multiple data sources and methods to develop the case study ensures that the issue is explored through a variety of lenses.
- **Flexibility (time):** case studies can be conducted in a few days, months or over several years and can be written up in different forms and lengths appropriate to the timescale. For instance, a case study on a specific P/CVE programme can focus on only some aspects of the programme (e.g. target groups, effectiveness, cost, etc.), and later be extended to provide an overview of other characteristics (e.g. transferability), although this may require additional data collection.
- **Flexibility (method):** case studies can draw upon whichever data/method is most helpful in understanding the case.

4.3.2. Why not?

Case studies also have a number of disadvantages, including:

- **Cost vs scope relationship:** case studies can be resource-intensive and expensive to carry out compared to the scope of their coverage, e.g. when conducting comparative case studies of several P/CVE programmes and relying on a number of data collection methods. Conducting a detailed, in-depth set of case studies spanning a range of different P/CVE areas/interventions is rarely feasible due to time and resource implications. Often a choice needs to be made between the depth (completeness and coverage) of individual case studies on the one hand, and the breadth (range/number) of case studies on the other. This choice reflects a trade-off between the depth of insights provided and the range of insights available.
- **Broader inferences:** as case study research typically focuses on a single case or a small group of cases, results usually cannot be generalised to the wider population. For instance, findings from a case study on a specific radicalised group are often not applicable to other radicalised groups.
- **Researcher bias:** case studies can be subject to the views and biases of individual researchers conducting them, since each case is a snapshot or reflection set within a particular context. Another source of possible bias is the limited range of perspectives gathered and data sources reviewed, which should be mitigated by appropriate triangulation of relevant data sets. To counter this, case studies should be chosen on a logical and transparent basis that reflects the purpose of the evaluation and (ideally) contributes to answering specific research questions.

- **Data processing:** the mass of data accumulated can sometimes be difficult to process, meaning that the richness of data can be lost when synthesising evidence on a particular case study. For example, when a P/CVE programme has been implemented in several locations and produced a large number of research outputs, it may be challenging to review all research outputs in great depth when producing a programme case study narrative.

4.4. Step-by-step application

A hallmark of case study research is the use of multiple data sources. Case study research can draw upon qualitative and/or quantitative methods. These may include, but are not limited to: literature review, interviews, observation techniques/ethnography, surveys and more. For a case study on the effectiveness of a counter-narrative campaign, this might typically include interviews with selected radicalised individuals and their families, a review of police records reporting on crimes committed by these individuals, and observation techniques.

The following is a step-by-step guide to how to design and apply case study research.

1. Define the research question
 - Identify an overarching research question (e.g. what are the success factors for a P/CVE intervention?). Is the aim to describe a case, explore a case, or compare between cases?
 - Focus on 'how', 'what' and 'why' sub-questions, for example how this P/CVE intervention was implemented, what the role of particular actors was in the implementation process, why specific actors were selected to be involved in the implementation process, etc.

2. Identify the case

- Identify the 'case' or unit of analysis, which may be a person, institution, programme, policy or system, for instance a radicalised group, a unit of a police force, a P/CVE programme, national policy for funding P/CVE programmes, etc.
- 'Bound' the case to ensure that the topic is within scope and relevant to overarching study objectives. Boundaries indicate what will and will not be studied in the scope of the research project. Note that cases can be bound by time, place, activity, definition and/or context. For instance, it can be agreed to only analyse P/CVE programmes operating in a specific country, operating for at least 3 years, only delivered online, etc.

3. Determine the case study type

- After determining the case and its boundaries, consider what type of case study will be conducted for what purpose, and which issues or questions it will help to address.
- The selection of a specific type of case study design will be guided by the timescale for the study and by the overall study purpose.
- Determine whether it is preferable to conduct a single case study (e.g. one radicalised group) or if a better understanding of the phenomenon will be gained through conducting multiple case studies (comparative case studies of a number of radicalised groups).

4. Select data collection methods

- Determine which methods (e.g. interviews, literature review, observation techniques, survey) are most appropriate for the case study in order to answer the research questions.

- Ensure data sources provide contrasting perspectives and can be used to test/reinforce each other.

5. Collect and analyse data

- Collect and analyse the data using appropriate analytical methods.
- If using multiple case studies, triangulate data, for instance by coding and scoring case study data to assist comparisons (e.g. comparing particular aspects of case studies against each other, such as cost, transferability, target groups, etc.), or by holding expert workshops (e.g. experts guiding comparative analysis of particular case studies).

6. Write up case study

- Write up case study narratives and cross-case analyses of coded and scored data.
- Extract recurring themes and explain links and contradictions across cases (e.g. by showing similarities in radicalised groups' recruitment strategies and differences in financing of groups).
- Assess, group and prioritise themes identified, and synthesis these findings in a case study narrative.

4.5. Ethics

When conducting case study research, the following ethical procedures should be adhered to (see also the ethics paragraphs in the methods used in case studies such as literature review, interviews, observational techniques/ethnography, and surveys):

- Clarify the purpose of the study and the anticipated audience for the information at the outset.
- Seek permission for access to documents, files and correspondence; these should

not be copied without explicit permission (e.g. police records, teachers' observations, etc.).

- Seek informed consent for each person interviewed, participant in a focus group and person observed.
- Conduct interviews on the principle of confidentiality and anonymity.
- Send a transcript (or note) to the interviewee and ask him/her to check it and give permission to use it for the research to ensure that the interviewee has given his/her consent. Typically, when conducting interviews with senior officials, the interviewer may be asked to send them a draft version of the interview transcript for review. Interviewees may give extra information and/or correct misunderstandings.
- Give interviewees the opportunity to see how their comments or observations are reported in the context of the case study and allow them to edit or add to their comments. For instance, the case study narrative can be shared with interviewees for review and the narrative adapted according to their comments and suggestions.
- Make sure not to report any data that a participant asks to be kept confidential, or confidential data from other sources (e.g. police reports).
- Note that direct attributable quotation and attributed judgements require the explicit permission of the respondent.
- Use pseudonyms in reporting individuals and institutions. While this does not guarantee anonymity, it reduces the likelihood that individuals and institutions will be identifiable.

List of references

Neumann, P. 2010. *Prisons and Terrorism: Radicalisation and De-radicalisation in 15 Countries*. London: The International Centre for the Study of Radicalisation and Political Violence.

Staun, J. (ed.). 2008. *Radicalisation, recruitment and the EU counter-radicalisation strategy*. COT Institute for Safety, Security and Crisis Management. As of 26 November 2014: <http://www.transnationalterrorism.eu/tekst/publications/WP4%20Del%207.pdf>

US General Accounting Office. 1990. *Case Study Evaluations*. As of 18 February 2018: http://betterevaluation.org/resources/guide/case_study_evaluations_gao

Check Appendix: Additional resources for further reading.



CHAPTER FIVE

Comparisons/benchmarking

Kei Ito and Anke van Gorp

See also:



Example(s) of use

The Institute for Strategic Dialogue's report 'European Counter-Radicalisation and De-radicalisation: A Comparative Evaluation of Approaches in the Netherlands, Sweden, Denmark and Germany' set out to examine different national strategies in the field of P/CVE. The study examined a large number of programmes from four EU Member States with diverging approaches and underlying challenges in countering the problem of radicalisation. The study did not use benchmarking as such, but employed a comparative approach to identify key challenges and lessons learnt from policies and interventions in these Member States. Although the report concluded that de-radicalisation programmes and policies are not simply transferable (as they will depend on the local context and culture), it is possible to identify good practice elements which can help inform policy and practices in the broad field of P/CVE (Butt and Tuck, 2014).

5.1. Brief description

Benchmarking is the method of comparing performance, processes or systems across and between countries, policy systems, policies, organisations and programmes. Benchmarking can also entail comparison of changes over time in one particular context. Whilst benchmarking originally started in private sector management, it has since been developed and is also used in the public

sector and by civil society organisations. It can be used to improve accountability and performance, to examine the allocation of resources and to identify and learn from best practice. It is a tool that can be used to identify what is being done, how and what outcomes are achieved by other interventions or organisations in the field of P/CVE that are seen as promising practices or high performers. This tool aids understanding of

how to do things better (more efficiently and/or more effectively) and what factors lead to that.

Benchmarking is a common evaluation method and a simple concept. However, there are often difficult decisions to make regarding what type of benchmarking to adopt (e.g. performance or process), what comparators to choose, what data to collect and how to analyse it, how to communicate the results, and how to make and monitor recommendations (Ling and Villalba van Dijk, 2009).

5.2. Purpose

There are numerous approaches to benchmarking, which can be grouped by their objectives. For example:

- *Performance benchmarking* is when the performance of a process is measured (with regard to outputs and outcomes) and compared against the performance of a similar process in a similar entity. In the context of a P/CVE intervention, benchmarking can be used to determine whether and how the results of an awareness-raising campaign among, for instance, families of foreign fighters, compare with those of similar P/CVE initiatives.
- *Process benchmarking* allows processes and procedures to be compared, even if their outputs and outcomes are different. For instance, process benchmarking may be used to identify success factors in one P/CVE intervention that may be transferable to another.
- *Domestic benchmarking* compares the process or performance of similar entities from within one country. This may include comparing the work or achievements of different not-for-profit organisations active in the field of P/CVE and operating in a

similar context, enabling them to learn from each other.

- *International benchmarking* compares the process or performance of similar entities from different countries. International benchmarking could, for instance, include comparing different countries' strategies and policies in tackling P/CVE at national level. It could also include comparing the work or achievements of enforcement agencies, although they would operate in different contexts.

5.3. When to use it

5.3.1. Why?

Benchmarking can be useful in a number of ways, including:

- To measure performance in a meaningful way by enabling comparisons;
- To generate ideas regarding what can be improved and how;
- To identify best practices and success factors of processes, interventions or policies in the field of P/CVE; and
- To share learning between parties with an interest in the compared entities (systems, policies, interventions, etc.).

5.3.2. Why not?

The use of benchmarking also has a number of disadvantages, including the following:

- It can be difficult to select a suitable comparator and access comparable and validated data relating to that comparator.
- Benchmarking needs to take into account that the context and some characteristics of compared organisations or interventions may differ and some measures or metrics used may differ across contexts.

- Benchmarking entities might be reluctant to share and pool data for comparison.

5.4. Step-by-step application

Generally, benchmarking follows a model of five stages as set out below (Ling and Villalba van Dijk, 2009).

1. Planning

The planning stage of a benchmark exercise involves identifying and selecting a suitable comparator or a number of comparators. This could include organisations or sub-units, P/CVE programmes, and de-radicalisation policies and/or policy systems that focus on similar issues or activities.

An important consideration when choosing a reliable comparator relates to data availability. It is useful to consider, for example, whether the data required is publicly available, whether the chosen comparator is willing to share the information, and whether the data available is reliable. Other important considerations are the potential limitations or weaknesses of the data (e.g. are recidivism rates in the comparators recorded similarly?).

The planning stage will also include data collection. The method for this will depend on what data need to be collected and who it should be collected from. For a benchmarking exercise of a P/CVE intervention, methods typically include document review, focus groups, interviews, surveys and field visits.

2. Analysis

There are several methods that may be used to analyse the data collected. The choice between those methods will depend on whether the data is qualitative (non-numerical, such as text or images) or

quantitative (numerical) in nature. Some key questions to help support the analysis are:

- What are the differences and similarities between the comparators?
- Is it possible to identify good P/CVE practices or bad P/CVE practices? What factors explain the success or failure of these examples?
- How does the overall context differ between the two comparators (for instance, in terms of institutional environment or local/national P/CVE policy framework)?
- What are the costs involved? Is one of the comparators more cost-effective?

3. Integration, action and monitoring

These stages of the benchmarking process are used to implement learning and follow up on the results of the previous steps. Whereas planning and analysis will lead to an indication of how one thing compares to the other, the integration, action and monitoring stages will allow findings to be communicated, recommendations to be turned into actions and monitoring plans to be devised in order to effectively respond to the outcomes of the benchmarking.

5.5. Ethics

In deciding what to benchmark (process or performance) and against which benchmark, the evaluator or the organisation commissioning the evaluation can influence what the result of the evaluation will be. If the evaluator has reason to think that the process went better than in another intervention he/she knows, then benchmarking against that intervention on process would lead to a positive result. This means that as long as there is no consensus about good practice, a

benchmark will always rely on some crucial subjective choices.

When benchmarking, evaluators not only need information about the intervention under evaluation but also about the benchmark, and this can lead to problems with regard to informed consent. If the evaluation of the benchmark was done under conditions of informed consent, did the respondents agree to re-use of their data as a benchmark? This might mean that in addition to obtaining informed consent from respondents in the intervention under evaluation, informed consent of respondents in the benchmark is also necessary. For the requirements on informed consent see for example the sections of this document on interviews, case studies and focus groups.

Ethical issues will also depend on the methods used to collect data, such as document reviews, focus groups, interviews, surveys and observational techniques.

List of references

- Butt, R. and Tuck, H. 2014. *European Counter-Radicalisation and De-radicalisation: A Comparative Evaluation of Approaches in the Netherlands, Sweden, Denmark and Germany*. Institute for Strategic Dialogue. As of 18 February 2018:
<https://www.counterextremism.org/resources/details/id/530/european-counter-radicalisation-and-de-radicalisation-a-comparative-evaluation-of-approaches-in-the-netherlands-sweden-denmark-and-germany>
- Ling, T. and Villalba van Dijk, L. 2009. *Performance Audit Handbook*. Santa Monica, CA: RAND Corporation.

Check Appendix: Additional resources for further reading.



CHAPTER SIX

Contribution analysis

Ben Baruch and Anke van Gorp

See also:



Example(s) of use

Contribution analysis was used to evaluate a wide ranging programme that included work on youth radicalisation in Burundi (EuropeAid, 2014). The evaluation focused on four areas: agricultural development, education, democratic governance (police, security, justice, media), and macroeconomic support. It covered a wide range of interventions over a long period of time and the objective was to identify the contribution of a selection of interventions to peace building.

6.1. Brief description

Contribution analysis is 'a form of theory-based evaluation where effectiveness is the main question and where experimental designs are not possible' (Wimbush et al. 2012, p.311). It is an established method of assessing the impact of interventions where project duration is short and evaluation resources are limited. It focuses on the steps along an intervention's theory of change (ToC), or logical framework (two terms that tend to be used interchangeably – please see the section on logic models for further explanation).

6.2. Purpose

The central aim of contribution analysis is to establish the validity of a project's/ programme's ToC/logical framework. It is one of the main approaches used not only to articulate but also to test programme theories. It seeks to provide a plausible explanation regarding the difference a programme is making to observed outcomes (Mayne, 2008).

6.3. When to use it

6.3.1. Why?

Contribution analysis is particularly useful when experimental or experimental evaluation

designs are difficult to implement, for instance due to issues with obtaining a counterfactual (i.e. what would have happened had the programme not existed (Wimbush et al. 2012, p.311)). Specific advantages of this method include the following:

- It can help evaluators 'come to informed arguments about the contribution of [an] intervention to observed conflict escalation or de-escalation' (Goldwyn and Chigas, 2013). In the context of P/CVE, experimental designs are particularly difficult to implement, for reasons of low resources, difficulties of obtaining agreements by site directors, or the low prevalence of P/CVE in the wider public, meaning impractically wide sampling is required.
- Rather than being a fully comparable alternative to experimental impact evaluation designs, Vaessen and Raimondo (2012) argued that contribution analysis should be considered a 'framework approach' that can encompass more specific methods to provide insight on elements of a 'causal intervention theory'. This flexibility of contribution analysis allows for a wider set of practical uses, even where budget, time and data are in short supply.
- It helps to develop a shared understanding of the 'contribution story' across stakeholders of and to articulate it in a coherent format that allows external groups to reach a clear understanding of the intended pathways through which outcomes are achieved.
- It benefits from the iterative nature of the process and the scope to draw on multiple sources of evidence.

6.3.2. Why not?

On the other hand, contribution analysis has the following disadvantages:

- It is tied to identified pathways and there is a risk that alternative and unexpected outcomes and routes to those outcomes may not be identified.
- Elements of the causal chain may remain weak despite this exercise, particularly where evidence is not available to support these elements and cannot be readily collected.
- It may fail to adequately account for feedback loops and wider contextual issues if not conducted carefully.

6.4. Step-by-step application

In practice, according to Mayne (2008), contribution analysis includes six steps:

1. Define the attribution problem

The first step of contribution analysis is to give some thought to the attribution problem, as conflict-prevention interventions typically generate questions about the extent to which the programme is responsible for observed results (Mayne, 2008). Once an attribution problem has been defined, a specific cause-effect question needs to be formulated. These range from the traditional ('To what extent has the programme caused the outcome?') to the managerial ('Is it reasonable to conclude that the programme has made a difference to the problem?') (Mayne, 2008:1).

2. Develop a theory of change/logical framework and ascertain risks to it

This step entails the building of a ToC/logical framework and a related results chain. These should identify

the assumptions underlying the chain's component parts and any risks to those assumptions, as well as any other factors that may shape outcomes (Mayne, 2008).

3. Collect available evidence on the theory of change/logical framework

The aim of this step is to begin assessing the underpinning process of the ToC/ logical framework. This approach can draw on various methods, information and data, ranging from reviewing existing evidence (past evaluations, monitoring data, etc.) through to conducting primary research and collecting data through interviews, surveys, etc.

4. Pull together and critically assess the contribution story

Based on the above information, the contribution story can be assembled. A contribution story is the narrative that explains how an intervention causes certain impacts. This version of the contribution story should be critically assessed to identify whether links are strong or weak, and for its overall credibility. In particular, the following questions should be considered (European Commission, 2013):

- 'Which links in the theory of change are strong (strong logic, good evidence available supporting the assumptions, low risk and wide acceptance) and which are weak?
- How credible is the story overall? Does the pattern of outcomes and links between them validate the contribution chain?
- Do stakeholders agree with the contribution story developed?
- Is it likely that any of the external significant factors have had a noteworthy influence on the results observed?

- What are the main weaknesses in the story? Where would additional data or information be useful?

5. Seek additional evidence

Based on the robustness assessment in step 4, this step entails obtaining additional evidence to improve the contribution story. Evaluators should specifically target the weaknesses identified in step 4, and consider what evidence could be obtained that will strengthen them. As in step 3, all impacts should be considered – both intended and unintended.

6. Revise and strengthen the contribution story

In this step, evaluators revise the contribution story, which may also include amendments to the ToC/logical framework. Depending on additional evidence, weaker parts of the earlier version will have either been given greater credibility or further discredited (or neither). The contribution story can be developed based on these findings, in line with the abovementioned emphasis on feedback loops.

The iterative process also requires that at this point evaluators consider whether to return to step 4 and re-examine the contribution story's strengths and weaknesses (Mayne, 2008). As Wimbush et al. (2012) note, the steps listed here do not end with the production of a refined contribution story, as there is still the matter of how the contribution story is used.

6.5. Ethics

When conducting a contribution analysis, the following ethical points should be considered:

- Ethical issues depending on the methods used to gather data – see for example interviews, surveys, focus groups, literature review, case studies, etc.

- The contribution analysis might be skewed depending on which stakeholders have been included. It may be in certain stakeholders' interests to emphasise certain parts of the contribution story. Even if stakeholders do not have such vested interests, the fact that they were involved in an intervention might convince them that it was a useful intervention because 'something was being done'. Social psychological research shows that being told how something should or could work strengthens one's belief that it does/did work. This confirmation bias might play a role if the evaluation is conducted ex-post (Roese and Vohs, 2012), and is particularly difficult to counter if those delivering an intervention are also evaluating it. Moreover, if stakeholders are already invested in a project, the 'sunk cost fallacy' might also cloud their judgment (Dobelli, 2013).⁵ This might compromise the accuracy of the contribution story.

List of references

Dobelli, R. 2013. *The art of thinking clearly*. London: Sceptre.

EuropeAid. 2014. *Joint evaluation of the cooperation with Burundi of Belgium, the European Commission, France, Germany, the Netherlands, Sweden and the United Kingdom*.

EuropeAid. As of: 13 August 2015: http://diplomatie.belgium.be/fr/binaries/Evaluation_Burundi_Executive_summary_EN_tcm313-247754.pdf

European Commission, Directorate-General for Regional and Urban Policies. 2013. *Evalsed Sourcebook: Method and Techniques*, European Commission. As of 18 February 2018:

http://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/evaluation_sourcebook.pdf

Goldwyn, R. and Chigas, D. 2013. 'Monitoring and evaluating conflict sensitivity: Methodological challenges and practical solutions'. As of 18 February 2018: <http://www.smallarmssurvey.org/fileadmin/docs/M-files/CCRVI/CCVRI-Monitoring-and-evaluating-conflict-sensitivity-challenges-and-solutions.pdf>

Mayne, J. 2008. *Contribution Analysis: An approach to exploring cause and effect. ILAC methodological brief. Institutional Learning and Change Initiative*. As of 18 February 2018: http://www.betterevaluation.org/en/resources/guides/contribution_analysis/ilac_brief

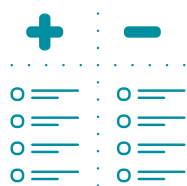
Roese, N. J. and Vohs, K.D. 2012. 'Hindsight Bias'. *Perspectives on Psychological Science*, 7(5):411–426.

Vaessen, J. and Raimondo, E. 2012. 'Making sense of impact: A methodological framework for assessing the impact of prizes'. *Evaluation*, 18:330–347.

Wimbush, E., Montague, S. and Mulherin, T., 'Applications of contribution analysis to outcome planning and impact evaluation', *Evaluation*, 18:310–329.

Check Appendix: Additional resources for further reading.

⁵ This is where people or organisations pursue an objective driven by how much effort has already been expended (sunk) rather than whether or not that objective is still worth pursuing.



CHAPTER SEVEN

Cost-benefit analysis

Benoit Guerin, Adam Martin and Anke van Gorp

See also:



7.1. Brief description

Cost-benefit analysis (CBA) is a form of economic evaluation used to assess the return on an investment made in a single (or group of) projects, programmes or policies. This requires an assessment of the costs and benefits (or outcomes) of an investment, where both are measured in monetary terms (Ling and Villalba van Dijk, 2009).

In common with cost-effectiveness analysis (CEA), which is an alternative form of economic evaluation, CBA may be undertaken to support decision makers in assessing a choice between two or more demands on resources within a given budget. In CBA, this would typically involve a comparison of their respective cost-benefit ratios (i.e. the monetary return per euro of investment). Alternatively, the decision on whether to fund a particular project, programme or policy might be made on the basis of whether it delivers a similar return on investment (i.e. cost-benefit ratio) to existing claims on resources.

As an example, CBA might be used to assess the return on investment in training courses

aimed at improving the responsiveness of first-line workers to early signs of radicalisation in terms of their cost-benefit ratio. However, in contrast to CEA, where benefits are measured in natural units (e.g. by improved knowledge of the practical skills required to recognise and deal with radicalisation), the key challenge associated with CBA is the need to monetise measured benefits. This can pose a significant challenge in terms of assessing the impact of interventions in policy areas (e.g. P/CVE) where the outcomes or benefits are not normally 'traded' in a free market, since a suitable 'price' or monetary value for those benefits may not be readily available. Furthermore, the process of monetising such benefits, or outcomes, could in some circumstances prove to be controversial in terms of public opinion.

Nevertheless, the key advantage of CBA (when compared to CEA or other forms of economic evaluation) is that it would be possible, at least theoretically, to draw comparisons between spending on P/CVE activities and other spending anywhere else in the public (or private) sectors of the economy. Thus, CBA may play a role in assessing whether or

not additional investment should be made in P/CVE activities overall, whereas it might be expected that CEA would be the preferred option for assessing competing demands on the existing P/CVE budget.

7.2. Purpose

CBA may be used to support comparisons of competing demands on national resources using monetary cost-benefit ratios. Thus, it might provide a useful tool for determining the impact (per euro spent) of P/CVE spending on the whole economy, rather than just in terms of impact of specific projects in the P/CVE sector. Assuming that the challenging process of monetising the benefits of P/CVE activities could be done in a credible way, such an analysis might form part of an argument for increased overall expenditure on P/CVE, and perhaps even for the transfer of resources from elsewhere in the public sector if it were deemed that P/CVE investment was associated with a higher cost-benefit ratio. In short, when compared to CEA, CBA can be used to demonstrate the ‘value’ that society places on investment in P/CVE activities when compared to other demands on resources – in other words, whether or not society considers the investment ‘worth it’.

In addition, CBA could also be used to provide decision makers with an indicator (i.e. the cost-benefit ratio in which costs and benefits are measured in monetary terms) of the relative value for money of specific courses of action, which can be used to inform judgements about which projects, programmes or policies to implement within a pre-determined P/CVE budget. However, if this is the sole purpose of the economic evaluation, CEA ought to be considered since this would not require the assignment of monetary value to benefits or outcomes.

7.3. When to use it

CBA may be used to provide an estimate of the return on investment of a project, programme or policy in terms of a monetary cost-benefit ratio. Such an assessment might form part of a trial or pilot study designed to ascertain the cost-benefit ratio associated with a new project, programme or policy before it is rolled-out more widely (ex-ante evaluation), or could be used to study a project, programme or policy that has already been implemented (ex-post evaluation). When there is uncertainty about the future costs and benefits of an intervention, an economic model may be developed. This could include, for example, a Markov or microsimulation model, such as a discrete event simulation model. However, CBA would likely only be used if there were sufficient, credible monetary values of both the costs and the benefits of an intervention from the perspectives of a wide range of different stakeholders and sectors of the economy.

7.3.1. Why?

CBA has the following advantages:

- It can support the investment of resources in new projects, policies or programmes which might otherwise be considered too controversial or lacking in sufficient evidence to be recommended by risk-averse policymakers;
- It can support decision makers in investing resources in existing projects that should be expanded but that might not otherwise be considered a candidate for further expenditure due to a lack of high-quality evidence on value for money;
- It can support decision makers in transferring resources away from existing projects which deliver relatively poor value for money (sometimes called ‘disinvestment’ or ‘divestment’) in order to

free up resources for investment in other projects which provide better value for money; and

- It can help decision makers in all policy areas demonstrate that they are delivering value for taxpayers' money. In many countries there is also a specific requirement to demonstrate improvements in productivity or year-on-year 'efficiency savings'.

7.3.2. Why not?

The use of CBA also has a number of drawbacks, particularly in the following settings:

- Where the monetisation of benefits is not necessary, for instance if comparisons are being made between projects with similar outcomes (e.g. increased knowledge of radicalisation processes amongst prison workers), meaning that CEA should be considered instead;
- Where insufficient resources (i.e. the total budget allocated to P/CVE activities) or significant capacity constraints (e.g. numbers of experts able to deliver interventions) mean that it is unrealistic to expect that a new project could ever be implemented, even if it has a favourable cost-benefit ratio;
- Where it is particularly challenging to gather sufficient, credible data on all the necessary costs and benefits of an intervention (e.g. if those costs and benefits occur a long time in the future), and hence CBA typically requires many assumptions;
- Where the various steps of a CBA (particularly the monetisation of benefits) require specific technical skills or input from specialists familiar with this methodology, meaning that it may be

considered prohibitively costly or time-consuming; and

- Where it is felt that political or other considerations will override any case for investment of government or EU resources on the basis of favourable cost-benefit ratios alone. The opposite can also be true in that wider considerations may not be fully considered because of the power of the economic argument provided by CBA.

7.4. Step-by-step application

The steps required to implement a CBA are outlined briefly below. These steps are adapted from various existing best-practice guidelines (e.g. Drummond et al. 2015; DG REGIO, 2013; HM Treasury, 2011). Whilst this section provides an overview of some of the issues, it is advisable to review more comprehensive guidance before embarking on a full economic evaluation.

1. Define the scope of the project and its objectives

It is critical that the scope of the evaluation and the research question are addressed at the outset. This is likely to include placing the evaluation in a particular decision-making context. Hence the following must be justified:

- *Time period included in the evaluation:* this could have a significant impact on the cost-benefit ratio. For example, some interventions could have a relatively favourable cost-benefit ratio in the short term, whereas others will present better value over the longer term.
- *Scope of the evaluation:* in terms of the sectors of the economy which should be included: for example, if the costs and benefits of a school-based intervention were measured only in terms of the costs and benefits to a school, then

this would likely produce very different estimates to the costs and benefits in terms of the rest of society. If the scope of the evaluation is too narrow, then it would be advisable to consider using CEA instead.

- *Feasible* (particularly from a logistical point of view): to estimate all costs and benefits related to the policy from the perspective of a wide range of different stakeholders across the whole economy.
2. Assess and quantify project costs
 - A full range of operating costs should be considered for inclusion in the analysis, including fixed costs (e.g. buildings) and variable costs (e.g. wages), capital costs (equipment, software, renting a venue, etc.), project management costs, staff costs, recurrent costs (material, supplies, printing, etc.), overheads (personnel training, etc.) and, depending on the perspective of the analysis, costs to the user (travel costs, etc.).
 - The analysis must consider 'opportunity costs' (or economic costs), rather than just 'unit costs' or 'market prices'. Opportunity costs are defined as the benefits that could have been obtained from the next best use of resources elsewhere in the economy, whereas unit costs or 'market prices' ought to be treated with caution because on their own they could distort the analysis and lead to misleading policy conclusions. For example, opportunity costs may differ from 'market prices' if volunteer labour is relied upon, or if the cost of buildings has already been paid from existing budgets. In these cases, the prices paid for those inputs in a new project will not correctly reflect the benefits that could be achieved if the resources had been used elsewhere.
 3. Assess and quantify project impacts or benefits
 - It is necessary to quantify costs in terms of their 'present value'. This means that if an expense was incurred several years ago, inflation should be taken into account to update the 'old' price to today's value (in other words, how much a particular P/CVE programme would have cost had it been implemented this year). Similarly, discounting should be applied to future costs, if only to reflect the savings to be made from postponing investment in terms of the potential interest that could be earned in the meantime (following standard guidelines such as those set out by HM Treasury (2011)). The following should be considered:
 - Allowance should be made for uncertainty in the estimates of costs using sensitivity analysis, or more advanced modelling techniques.
 - Economies (or diseconomies) of scale may impact on the average cost of delivering a project. For example, average costs could be lower if the scheme is rolled out across the whole country when compared to running a smaller pilot scheme.
 - For a full CBA, the evaluation should ideally include direct benefits (such as increased knowledge of the radicalisation process amongst school teachers) and indirect, downstream benefits (such as reduced benefits payments by the state, fewer crimes committed, etc.), as well as intermediate (e.g. number of young people engaged in risky behaviours in juvenile crime statistics) and final (e.g. number of prisoners sentenced for committing a P/CVE-related crime) outcomes.

- The key challenge in CBA is the necessity of placing a monetary value on those benefits of a project which cannot be readily determined from (free-) market data (for example the monetary value of a reduction in the numbers of individuals becoming radicalised and involved in radicalised crimes). This process could involve a 'revealed preference' approach which infers a price from observing consumer behaviour, or a 'willingness to pay' or 'stated preference' approach which can be used to estimate values by asking people what they would be willing to pay for a particular benefit. This can be achieved using a 'discrete choice model', for example, which is a form of virtual experiment (see HM Treasury, 2011).
 - The gold-standard method for assessing the impact of a policy is the randomised control trial (RCT) in which units (e.g. people, places) are randomised to intervention and control groups. In the absence of a RCT, results should be treated with caution because of the risk of endogeneity or self-selection bias; however, a range of econometric or quasi-experimental methods (including difference-in-difference, instrumental variables and fixed-effects models) may provide a promising alternative in policy areas where RCTs are not feasible (Martin, 2014).
 - Consistent with the estimation of costs, the future impact of projects should be discounted at an appropriate rate, and allowance made for uncertainty in the parameter estimates.
4. Calculate the cost-benefit ratio
- The cost-benefit ratio is used to express the results of the analysis. Put simply,

it is the ratio of additional (or marginal) costs of a new project to the additional (or marginal) benefits of the project, expressed in monetary terms.

5. Select or amend the policy, programme or intervention

Competing interventions can be ranked in terms of their value for money using the cost-benefit ratio in order to provide a decision maker with a 'menu' of alternative options, with due consideration of the limitations of the study design (i.e. the internal validity) and applicability of the study to the policy decision under consideration (i.e. the external validity).

7.5. Ethics

When conducting cost-benefit analyses, the following ethical points should be considered:

- If only a limited range of costs or benefits were included in the evaluation (e.g. if the evaluation considered only the short-term cost-effectiveness from the perspective of local government) then policy decisions might be based on evidence which has overlooked some important potential winners and losers of those decisions.
- Placing a value on social benefits or attempting to systematically quantify benefits can be problematic since it may ignore some of the underlying complexity.
- People might claim that some things cannot be monetised. These may include, for example, a reduction in crime, but also improvements in human life, health, happiness, wellbeing, safety or the natural environment. Even if these aspects can be monetised, the value placed on each may

not be shared by all. (In the case of crime, however, there are unit costs available.)⁶

- Obtaining data on certain aspects of the programme may generate privacy issues (e.g. in relation to wages, effects of P/CVE intervention on political or religious convictions, mental health, etc.).
- There has been some debate about whether or not it is ethical to 'deny' citizens access to effective interventions on the basis of that they are not cost-effective. For an example of this argument in the healthcare setting, see Harris (2005).
- Arguments based on ideas of distributive justice or the equity–efficiency trade-off might be overlooked in CBA. For instance, it may be necessary to incorporate equity considerations relating to income to ensure that interventions are allocated 'fairly' across society regardless of their relative cost-effectiveness.
- People generally prefer interventions that do not restrict individual choice or impact unnecessarily on civil liberties. In some cases, it is possible that these considerations would override a concern for choosing the intervention with the most favourable cost-benefit ratio. An intervention with a slightly less favourable cost-benefit ratio but with less impact on civil liberties might therefore be preferable. This also depends on what costs are taken into account – an infringement on civil liberties or a limitation on individual choice might be difficult to 'price' and therefore might be left out of a cost-benefit analyses.

List of references

- Drummond, M.F., Sculpher, M.J., Claxton, K., Stoddart, G.L. and Torrance, G.W. 2015. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press.
- European Commission, Directorate-General for Regional and Urban Policies. 2013. *Evalsed Sourcebook: Methods and Techniques*. European Commission.
- Harris, J. 2005. 'It's not NICE to discriminate'. *Journal of Medical Ethics*, 31(7):373–5.
- HM Treasury. 2011. *The Green Book: Appraisal and Evaluation in Central Government*. London: HM Treasury.
- Ling, T. and Villalba van Dijk, L. 2009. *Performance Audit Handbook*. Santa Monica, CA: RAND Corporation.
- Martin, A., Ogilvie, D. and Suhrcke, M. 2014. 'Evaluating causal relationships between urban built environment characteristics and obesity: a methodological review of observational studies'. *International Journal of Behavioral Nutrition and Physical Activity*, 11:1–15.

Check Appendix: Additional resources for further reading.



CHAPTER EIGHT

Cost-effectiveness analysis

Benoit Guerin, Adam Martin and Anke van Gorp

See also:



8.1. Brief description

Cost-effectiveness analysis is a form of economic evaluation used to inform the choices faced by decision makers when allocating a given budget to competing demands. It seeks to promote value for money, whereby maximum impact is achieved for minimum cost, by comparing two or more projects, programmes or policies in terms of both their costs and their consequences. For example, CEA could be used to compare different training courses aimed at improving the responsiveness of frontline workers to early signs of radicalisation in terms of their costs (measured in monetary units) and effectiveness (e.g. measured by improved knowledge of the practical skills required to recognise and deal with radicalisation).

The results of a CEA are typically reported as a ratio of the marginal (or incremental) costs of a particular intervention (usually a proposed new intervention) divided by the marginal benefits when compared to some alternative course of action (usually current practice). A list of competing interventions ranked according to

their relative cost-effectiveness may then be presented to decision makers. Alternatively, the decision about whether to fund a particular project, programme or policy can be made on the basis of whether it delivers a similar or improved cost-effectiveness ratio when compared to existing claims on resources.

8.2. Purpose

CEA is suited to the comparison of alternative courses of action within a specific sector of the economy or policy area that have comparable objectives which are not easily expressed in monetary terms (e.g. increased knowledge of radicalisation processes among prison workers). This is in contrast to CBA, which supports comparisons of competing demands across all sectors of the economy using monetary cost-benefit ratios.

CEA provides decision makers with an indicator (i.e. a cost-effectiveness ratio with a common denominator) of the relative cost-effectiveness of specific courses of action which can be used to inform judgements about which projects, programmes or policies to implement within

a pre-determined P/CVE budget. For instance, in order to improve outcomes for a given cost, a CEA might be used to make a case for transferring resources from one project to another (e.g. from a project focused on the prevention of radicalisation in secondary schools to another project focused on primary schools, should the latter prove to be more cost-effective). This is in contrast to CBA, which might be more appropriate for making an argument for increased overall expenditure on P/CVE by drawing comparisons between the monetary return on investment when compared to other national spending (e.g. defence spending).

8.3. When to use it

Cost-effectiveness analysis (CEA) may be used to evaluate the costs and consequences of comparable projects in which consequences are assessed using a single, non-monetised unit of measurement (e.g. number of police officers trained to recognise and identify vulnerable individuals who may be at risk of radicalisation). It might form part of a trial or pilot study designed to ascertain the cost-effectiveness of a new project, programme or policy before it is rolled out more widely (ex-ante evaluation), or could be used to study an existing project, programme or policy that has already been implemented (ex-post evaluation). When there is uncertainty about the future costs and benefits of an intervention, an economic model may be developed. This could include, for example, a Markov or microsimulation model, such as a discrete event simulation model. However, CEA can only be applied when there is sufficient data (from primary or secondary data sources or an economic model) on (at least) the relative costs and consequences of an intervention when compared to a credible alternative course of action, before and after the intervention took place.

8.3.1. Why?

CEA can be advantageous when used for the following:

- To enable the costs and benefits of competing programmes to be compared in such a way that may not be possible using standard CBA techniques because of significant challenges in monetising or valuing outcomes;
- To ensure that a given P/CVE budget has maximum impact by supporting decision makers in making judgements about competing projects, programmes or policies based on their relative cost-effectiveness (e.g. by comparing a three-month anti-radicalisation programme to a six-month programme which, for instance, might be more costly and only slightly more effective);
- To support the investment of resources in new projects, policies or programmes which might otherwise be considered too controversial or lacking in sufficient evidence to be recommended by risk-averse policymakers;
- To support decision makers in investing resources in existing projects that should be expanded but that might not otherwise be considered a candidate for further expenditure due to a lack of high-quality evidence on value for money;
- To support decision makers in transferring resources away from existing projects which deliver relatively poor value for money (sometimes called 'disinvestment' or 'divestment') in order to free up resources for investment in other, more cost-effective projects;
- To help decision makers in all policy areas to demonstrate that they are delivering value for taxpayers' money. In many countries there is also a specific requirement to

demonstrate improvements in productivity or year-on-year 'efficiency savings'.

8.3.2. Why not?

CEA also has a number of drawbacks, particularly in the following settings:

- When insufficient resources (i.e. the total budget allocated to P/CVE activities) or significant capacity constraints (e.g. numbers of experts able to deliver interventions) mean that it is unrealistic to expect that a new project could ever be implemented, even if it is shown to be cost-effective. In this case, despite being more technically challenging since it would be necessary to monetise the benefits, CBA may be better suited to making an argument for increased resources being allocated to P/CVE activities overall.
- Where it is particularly challenging to gather sufficient, credible data on all the necessary costs and benefits of an intervention and a suitable comparator (particularly if those costs and benefits fall a long time in the future, for example, or across a wide range of different stakeholders or sectors of the economy). For example, there may be a particularly complex project that is designed to deliver multiple objectives which cannot be fully captured in an evaluation (European Commission, 2013). It may also require many assumptions.
- Where the various steps of a cost-effectiveness analysis require specific technical skills or input from specialists familiar with this methodology. This might be considered to be prohibitively costly or time-consuming.
- Where it is felt that political or other legitimate considerations will override any case for investment of government or EU resources on the basis of cost-effectiveness evidence alone. The

opposite can also be true in that wider considerations are not fully considered because of the power of the economic argument provided by this method.

8.4. Step-by-step application

The steps required to implement a CEA are outlined briefly below. These steps are adapted from various existing best-practice guidelines (e.g. European Commission, 2013; Drummond et al. 2015). Whilst this section provides an overview of some of the issues, it is advisable to review more comprehensive guidance before embarking on a full economic evaluation.

1. Define the scope of the project and its objectives

It is critical that the scope of the evaluation and the research question are addressed at the outset. This involves determining what the main objectives of the project are. The purpose here is to ensure:

- That the basic requirements of a CEA are incorporated into the study (i.e. that both costs and consequences are measured for an intervention group and control group before and after the intervention was implemented);
- That the evaluation is placed in a particular decision-making context. Hence the following must be justified:
 - The choice of comparison group, including whether or not a 'do nothing' alternative was considered. For instance, a new intervention to support families of children at risk of radicalisation in their own homes might appear highly cost-effective when compared to an existing home-based scheme which is very expensive and performing poorly, whereas it may in fact deliver poor value for money

when compared to another school-based or 'do nothing' approach.

- Why some relevant alternative comparison groups may have been omitted.
- The perspective (or viewpoint) of the evaluation, which would likely have a significant impact on the costs and impacts to be incorporated in the analysis. The perspective could be that of secondary schools, prisons, the Home Office, the Department of Education, the EU, individuals at risk of radicalisation, the wider local community, the national or international economy, etc.).
- The time period included in the evaluation, which could have a significant impact on the relative cost-effectiveness of interventions (e.g. some interventions may appear relatively cost-effective in the short term, whereas others will demonstrate better value over the longer term).

2. Assess and quantify project costs

- A full range of operating costs should be considered for inclusion in the analysis, including fixed costs (e.g. buildings) and variable costs (e.g. wages), capital costs (equipment, software, renting a venue, etc.), project management costs, staff costs, recurrent costs (material, supplies, printing, etc.), overheads (personnel training, etc.), and, depending on the perspective of the analysis, costs to the user (travel costs, etc.).
- The analysis must consider 'opportunity costs' (or economic costs), rather than just 'unit costs' or 'market prices'. Opportunity costs are defined as the benefits that could have been obtained from the next best use of resources elsewhere in the

economy, whereas unit costs or 'market prices' ought to be treated with caution because on their own they could distort the analysis and lead to misleading policy conclusions. For example, opportunity costs may differ from 'market prices' if volunteer labour is relied upon, or if the cost of buildings has already been paid for from existing budgets. In these cases, the prices paid for those inputs in a new project will not correctly reflect the benefits that could be achieved if the resources had been used elsewhere.

- It is necessary to quantify costs in terms of their 'present value'. This means that if an expense was incurred a few years ago, inflation should be taken into account to update the 'old' price to today's value (in other words, how much a particular P/CVE programme would have cost had it been implemented this year). Similarly, discounting should be applied to future costs, if only to reflect the savings to be made from postponing investment in terms of the potential interest that could be earned in the meantime (following standard guidelines such as those set out by HM Treasury (2011)).
 - Allowance should be made for uncertainty in the estimates of costs using sensitivity analysis, or more advanced modelling techniques.
 - Economies (or diseconomies) of scale may impact on the average cost of delivering a project. For example, average costs could be lower if the scheme is rolled out across the whole country when compared to running a smaller pilot scheme.
- ## 3. Assess and quantify project impacts or benefits
- Depending on the perspective of the analysis, the evaluation could include

direct benefits (such as increased knowledge of the radicalisation process amongst school teachers) or indirect, downstream benefits (e.g. reduced benefits payments, fewer crimes committed, etc.), and intermediate (e.g. number of young people engaged in risky behaviours in juvenile crime statistics) or final (e.g. number of prisoners sentenced for committing a P/CVE-related crime) outcomes. Nevertheless, a core feature of the CEA is measurement of impact in a single, universal outcome measure.

- The impact or benefits should typically be based on primary or secondary data, or economic modelling.
 - Consistent with the estimation of costs, the future impact (or benefits) of projects should be discounted at an appropriate rate, and allowance made for uncertainty in the parameter estimates.
4. Calculate the incremental cost-effectiveness ratio
- The incremental cost-effectiveness ratio is used to express the results of the analysis. It expresses the incremental (or marginal) costs and benefits of a particular P/CVE programme or intervention in relation to an appropriate alternative course of action (typically usual practice).
5. Select or amend the policy, programme or intervention
- Competing interventions can be ranked in terms of their value for money using the cost-effectiveness ratio in order to provide a decision maker with a 'menu' of alternative options (Better Evaluation, 2014).
 - It is important to note that best value for money does not necessarily mean the lowest-cost option (i.e. the most cost-effective P/CVE programme is not

necessarily the programme with the lowest cost).

- The results of a CEA should be considered in the context of other existing, comparable cost-effectiveness evidence (if available), with due consideration of the limitations of the study design (i.e. the internal validity) and applicability of the study to the policy decision under consideration (i.e. the external validity).

8.5. Ethics

When conducting a CEA, the following ethical points should be considered:

- Since the purpose of CEA is to assess the potential for alternative interventions to deliver a given outcome (or policy objective), no attempt is made at 'valuing' the outcome measure, and so it must be implicitly assumed that the objectives are 'worth having' from the perspective of individuals or society. In order to understand the 'value' of a given outcome or policy objective (e.g. a reduction in people engaging with extremists online), it may instead be appropriate to undertake a CBA and ascertain the monetary 'value' of a particular outcome or policy objective. This may be done through a contingent valuation or revealed preference study, for example, although this would inevitably pose significant challenges.
- If only a limited range of costs or benefits are included in the evaluation (e.g. if the evaluation considered only the short-term cost-effectiveness in relation to the perspective of local government) some important winners and losers of a proposed policy decision may have been overlooked.
- It may be desirable to include 'social value judgments' in the CEA. For instance, it may

be necessary to incorporate important equity considerations relating to income, to ensure that interventions are allocated 'fairly' across society regardless of their relative cost-effectiveness. In the healthcare setting, for example, societies generally have a preference for tackling inequality in health and hence may favour interventions which deliver a little less health overall, but which ensure that health is more evenly distributed (this is sometimes known as the equity–efficiency trade-off).

- Obtaining data on certain aspects of the programme may generate privacy issues (e.g. in relation to wages, effects of P/CVE intervention on political or religious convictions, mental health, etc.).
- There has been some debate about whether or not it is ethical to 'deny' citizens access to effective interventions on the basis of that they are not cost-effective (for an example of this argument in the healthcare setting, see Harris (2005)).
- People generally prefer interventions that do not restrict individual choice or impact unnecessarily on civil liberties. In some cases, it is possible that these considerations would override a concern for choosing the most cost-effective intervention. An intervention with a slightly less favourable cost-effectiveness ratio but with less impact on civil liberties might therefore be preferable. This also depends on what costs are taken into account – an infringement on civil liberties or a limitation on individual choice might be difficult to 'price' and might therefore be left out of a cost-effectiveness analyses.

List of references

Better Evaluation. 2014. 'Cost Effectiveness Analysis'. As of 18 February 2018: <http://betterevaluation.org/evaluation-options/CostEffectivenessAnalysis>

Drummond, M.F., Sculpher, M.J., Claxton, K., Stoddart, G.L. and Torrance, G.W. 2015. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press.

European Commission, Directorate-General for Regional and Urban Policies. 2013. *Evalsed Sourcebook: Methods and Techniques*. European Commission.

HM Treasury. 2011. *The Green Book: Appraisal and Evaluation in Central Government*. London: HM Treasury.

Martin, A., Ogilvie, D. and Suhrcke, M. 2014. 'Evaluating causal relationships between urban built environment characteristics and obesity: a methodological review of observational studies'. *International Journal of Behavioral Nutrition and Physical Activity*, 11:1–15.

Check Appendix: Additional resources for further reading.



See also:



CHAPTER NINE

Cross-sectional study and data analysis

Megan Sim and Anke van Gorp

Example(s) of use

The Violence Prevention Network⁷ is implemented in several states in Germany. The programme used cross-sectional data analysis to understand the effectiveness of its work in terms of levels of recidivism (Lukas, 2012). After the programme (though the point in time is not stated) recidivism (the outcome variable) was lower among those who participated in the programme than among those who did not.

A new method of assessing vulnerability to violent radicalisation was used in a study to identify areas of focus for preventive intervention. The methods included a cross-sectional survey of a representative population sample of men and women aged 18–45, of Muslim heritage and recruited by quota sampling by age, gender and working status, in two English cities. The main outcomes include self-reported health, symptoms of anxiety and depression, and vulnerability to violent radicalisation as assessed by sympathies for violent protest and terrorist acts (Bhui et al. 2014).

9.1. Brief description

A cross-sectional approach to evaluation focuses on one point in time (as opposed to conducting the evaluation over an extended period of time), and as such, cross-sectional

data analysis refers to the analysis of data that has been collected at one point in time (Liu, 2008). For the purpose of evaluation, cross-sectional data analysis can be used to get a 'snapshot' of the group for which data has

7

<http://www.violence-prevention-network.de/en/>

been gathered. The general approach of cross-sectional data analysis involves analysing the relationship between one or more independent variables (also known as explanatory variables, which explain the variation in the value of the dependent variable) and a dependent variable (also known as an outcome variable).

The term cross-sectional is often used to describe questionnaire or survey studies (Liu, 2008); for instance, the evaluator may distribute a survey to participants of a P/CVE intervention to collect information about their personality and satisfaction with the intervention approach. The data can then be analysed to determine whether participants' personalities are associated with their satisfaction. However, cross-sectional data need not always be collected via questionnaires and may be drawn from other sources such as administrative documents (Liu, 2008).

9.2. Purpose

One objective of cross-sectional data analysis is to assess the relationship between independent variables to a dependent variable (Hall, 2008). For instance, in a cross-sectional study involving participants of a P/CVE intervention, an analysis can examine the associations between different demographic variables and the effectiveness of the intervention.

Cross-sectional data analysis may also be used to make inferences about a wider population if the participants included in the analysis have been randomly sampled. Where random sampling has been used, it is assumed that the group being studied is representative of the population of interest – in other words, that the characteristics of the group (e.g. demographic factors, behavioural factors) reflect the larger population from which it is drawn (Treiman, 2009). For instance, adult right-wing extremists who have been randomly chosen to participate in a P/CVE intervention may be assumed to be representative of the wider population

of adult right-wing extremists, and thus the findings of cross-sectional analyses on the intervention group may be generalised to the wider population. This assumption is important for drawing broader conclusions about the population from the analysis. It is possible to conduct cross-sectional data analyses with a sample that has not been randomly selected (and therefore is not assumed to be representative of the wider population), though the findings of this analysis should not be generalised beyond the sample.

It is worth noting that cross-sectional data analysis can help establish the significant relations between certain variables and factors, but in the absence of a randomised trial (or strong instrumental variable) cannot determine causation (Liu, 2008). This is due to the numerous variables influencing any outcome, and the difficulty of tracing the impact of a single factor or variable.

9.3. When to use it

Cross-sectional data analysis can be used to examine the relationship between one or more independent variables and a dependent variable, and can be used at any point in the evaluation when data on the relevant independent and dependent variables are available.

9.3.1. Why?

Cross-sectional approaches have the following advantages:

- Data collection is relatively straightforward. For instance, if using surveys:
 - Participants are usually more willing to cooperate in a one-off survey than a series of surveys.
 - Compared to longitudinal or panel surveys, cross-sectional data are less expensive and time-consuming to gather.

- Data that is already being collected (e.g. administrative data, census data) can be used for other purposes.
- They are highly efficient when used in an exploratory way to determine if there are relationships between variables.
- They can be used to test a model that has been proposed on a theoretical basis.

9.3.2. Why not?

The disadvantages of cross-sectional approaches include the following:

- They do not allow causal effects to be established. For instance, an evaluation might find that participants who believed that a P/CVE intervention was effective also expressed satisfaction with the intervention. However, based on this association it would not be possible to determine whether participants were satisfied because they believed the intervention was effective, or whether they believed the intervention was effective because they were satisfied with it.
- They do not make it possible to determine whether there are additional variables that account for the observed relations. For instance, participants might be satisfied with a P/CVE intervention and believe that it is effective because of variables that have not been accounted for, such as the reputation of the organisation carrying out the intervention.
- They do not allow for changes to be examined over a period of time. Cross-sectional data analysis captures a snapshot and is not appropriate for examining trends over time.
- Depending on the method used to collect the cross-sectional data (e.g. surveys or interviews) there is potential for the data to be affected by respondent or interviewer

bias (see for example Chapters 14 and 27 for more information).

9.4. Step-by-step application

The steps involved in a cross-sectional approach are as follows:

1. Determine the variables of interest to the evaluation
 - This includes the dependent (or outcome) variable and the independent variables (or explanatory variables).
 - The dependent variable chosen will most likely be linked to the purpose of the evaluation. For instance, a process evaluation might consider dependent variables such as participants' satisfaction with a P/CVE intervention, while an impact evaluation might consider dependent variables such as the effectiveness or cost-effectiveness of a P/CVE intervention.
 - The choice of independent variables may be linked to the theoretical basis of the intervention. For instance, an evaluation of a P/CVE intervention that targets participants' cognition may include independent variables such as participants' self-reported identity, emotions and beliefs.
2. Collect data
 - Data can be collected in a number of ways, for instance via a questionnaire or survey, or by retrieving administrative or other existing data.
 - Data can be inputted into a spreadsheet in preparation for data analysis.
3. Conduct cross-sectional data analysis
 - Data can be analysed in a software package such as SPSS, Stata or R.
 - Cross-sectional data analysis generally takes the form of bivariate analysis

(analysis with two variables) or multivariate analysis (analysis with more than two variables).

- Bivariate analysis includes statistical tests such as Pearson's Chi-square, Student's t-test, one-way analysis of variance and Pearson's correlation.
- Multivariate analysis includes statistical models such as linear regression, logistic regression, and analysis of variance.
- The results of the data analysis can be used to inform the evaluation of the P/CVE intervention or programme. Interpretation of the data analysis should be considered in the context of certain caveats. For instance, in the absence of a randomised trial or strong instrumental variable, causal effects cannot be inferred and the generalisability of the findings may be limited depending on the sample from which the data has been collected.

9.5. Ethics

The ethical issues in cross-sectional data analysis depend on the method(s) of gathering data. When existing personal data are used in a cross-sectional data analysis, evaluators need to adhere to data protection regulations. This means that personal data can only be used for the purpose for which they were collected (this is known as purpose binding). Personal data collected for another purpose (for example the implementation of social policy) cannot simply be used for an evaluation unless respondents were informed that the data would be used in that manner.

If surveys, questionnaires or interviews are used to collect data, two important ethical issues that need to be addressed are:

- Confidentiality: the respondent's right to confidentiality should always be respected and any legal requirements on data protection should be adhered to.

- Informed consent: respondents should be fully advised on the aims of the survey so that they may make an informed judgment about whether they wish to participate. Respondents' consent to participate in the survey must be obtained and recorded. However, it is not necessary to do this in a separate step. It may be sufficient to include a sentence about this when contacting the respondent (either prior to filling the survey in the case of a self-completion survey, or at the time of taking the survey in the case of an interview-based survey). If the respondent proceeds to take the survey, this can be taken to represent implicit consent to the use of their data.

List of references

- Bhui, K., Warfa, N., and Jones, E. 2014. 'Is violent radicalisation associated with poverty, migration, poor self-reported health and common mental disorders?'. *PLoS one*, 9(3):e90718.
- Liu, C. 2008. 'Cross-sectional data' in Lavrakas, P.J. (ed.) *Encyclopedia of Survey Research Methods*, pp. 171–173. Thousand Oaks, CA: SAGE Publications, Inc.
- Lucas, H. 2012. Untersuchung zur Legalbewährung der Teilnehmer an VPN-Trainingskursen im Jugendstrafvollzug Kommentierte Kurzfassung, Berlin: Violence Prevention Network.
- Hall, J. 2008. 'Cross-sectional survey design' in Lavrakas, P. J. (ed.) *Encyclopedia of Survey Research Methods*, pp. 173–174. Thousand Oaks, CA: SAGE Publications, Inc.
- Treiman, D. J. 2009. *Quantitative data analysis*. San Francisco, CA: John Wiley and Sons, Inc.

Check Appendix: Additional resources for further reading.



See also:



CHAPTER TEN

Data mining

Sarah Grand-Clement and Anke van Gorp

Example(s) of use

Defence agencies around the world use data mining to look for threats around the web. These agencies include the US National Security Agency, which has used programmes such as Global Clarity and Prism, and the UK's Government Communications Headquarters, which uses Tempora (MacAskill et al. 2013).

Indeed, data mining has become a growing method when looking for threats online. Also part of the US National Intelligence office is the Intelligence Advanced Research Projects Activity (IARPA). One of their research programmes is Open Source Indicators, which uses data mining to 'anticipate and/or detect significant societal events, such as political crises, humanitarian crises, mass violence, riots, mass migrations, disease outbreaks, economic instability, resource shortages, and responses to natural disasters' (IARPA, n.d.).

10.1. Brief description

Data mining is a technique used to sort through large amounts of data and extract patterns, relationships and other useful information. It is part of a process that leads to what is known as Knowledge Discovery in Databases, which is 'the overall process of discovering knowledge from data' (Abonyi and Feil, 2000, p.x). This process includes data selection, data cleaning and data reduction, all leading to data

mining, which is followed by the evaluation and interpretation of these patterns (Fayyad et al. 1996).

Data mining includes several different techniques that show different types of patterns and relationships within the data. These may include:

- **Clustering:** this technique groups together data (about people, objects, websites, etc.) based on their levels of similarity,

thus creating clusters. This can be used to extract social networks. This then helps with crime analysis, for example, where the use of clustering can identify high-risk areas.

- **Classification:** this technique divides data by class. As stated by Saiya and Scime (2014, p.2), the classification technique shows 'the precise ways in which explanatory variables work together to show which countries are especially prone or immune to terrorism, when and why'. For example, if searching through websites relating to violence, this technique would subdivide the websites into different categories, such as physical and emotional violence, with both categories then having further offshoots such as guns, blackmail or chemical weapons.
- **Association:** this technique, also known as the 'if/then' method, is used to uncover relationships between events happening in the same timeframe. For example, this can be practical when analysing client preferences and finding associations in which websites are visited by clients before they engage in violent action.

Various graphics are associated with each technique, allowing the data to be visualised in a simpler way.

10.2. Purpose

Data mining can be used as a first step when presented with a lot of data (such as from Twitter, or high-volume administrative data). For example, if evaluators want to analyse websites that may incite violence, data mining could offer tools to extract useful information that would be grouped in logical categories chosen by the evaluators. For instance, these could be grouped into websites inciting violent action through propaganda, websites that raise funds for violent groups, and websites that are used to build a community of people involved

in violent extremism. The use of various data-mining methods also shows the relationship between these various websites.

10.3. When to use it

10.3.1. Why?

The advantages of data mining include the following:

- Data mining not only highlights which variables are important but makes it possible to identify patterns and extract useful knowledge from large amounts of data which otherwise would not yield any meaningful information. It is supported by a very large amount of evidence or different sources. Data mining can be used in conjunction with regression to predict future outcomes (see e.g. Oswald et al. 2018). However, this is based on the behaviour staying the same and following the same pattern.
- Data mining can also be used to evaluate behaviour after an intervention, and detecting whether this has been successful and has brought about a change.

10.3.2. Why not?

There are also disadvantages to the use of data mining, namely:

- There are risks around the use of personal information that is not made public, loss of information related to data mining, and the way in which the information is collected and used. There have recently been security breaches involving national security agencies, highlighting the importance of taking into account ethical and legal considerations when collecting and using data.
- Data may not always be accurate. The analysis is intrinsically dependent on the

quality and completeness of the dataset and is also subject to its inherent biases. With large volumes of data it is not always possible to review for quality and bias, so it is necessary to be aware of them in the analysis. In addition, associations in the data may be correlated but not causally related. Moreover, there is the possibility that information could be classified inaccurately.

- Data mining does not provide information on how people think and act, or about cultural norms that impact on how local social networks are formed. It should not be used as a replacement for understanding cultural or social behaviour.

10.4. Step-by-step application

Data mining involves the following steps:

1. Identify the research aim of the data mining
 - While data mining helps in identifying patterns and connections, there needs to be a system in place beforehand to collect the data and prepare it.
2. Select a search method
 - There are two main ways in which information can be searched in the context of tackling radical violence:
 - Pattern-based data mining: this requires a 'feedback mechanism' in order to help the statistical package undertaking the analysis to recognise relevant information automatically.
 - Subject-based data mining: this focuses the search for information on finding relationships between individuals within the data for a more focused analysis.
3. Analyse mined data
 - Making use of 'information fusion', which is used to group together data from

various different sources, is useful in order to make more accurate inferences. This technique is particularly useful within P/CVE as it makes it possible to compile data concerning travels, monetary transactions, meeting details, etc.

- Identifying relationships and patterns via data mining takes place through the use of statistical packages such as SPSS. With the correct information inputted within the system, this then creates visual maps such as clusters or classification.

10.5. Ethics

There are several ethical issues in data mining depending on the type of data used (personal e.g. personal or non-personal data), whether data from different databases are combined and what conclusions are drawn. For example:

- If personal data are used in data mining, evaluators need to adhere to data protection regulations. This means that personal data can only be used for the purpose for which they were collected (purpose binding). Personal data collected for another purpose (for example the implementation of social policy) cannot simply be used for an evaluation unless respondents were informed that the data would be used in that manner.
- As stated above, data might not be accurate. Correcting data is difficult, labour-intensive and sometimes impossible because data are shared with so many other databases that it becomes impossible to correct a flaw in every database.
- Even if collecting data from public sources, attempting to identify individuals from that data may violate local data protection laws. For example see the controversy around

the 'unmasking' of the artist Banksy, which was based on research by Hague et al. (2016). Note however, the method used does have applications in P/CVE, as noted by the authors.

- People might not know that data are gathered about them and shared with others, and/or there might not be informed consent to do so, which is necessary if personal data are used.
- If the data used are not personal data (i.e. do not include any information that can lead to the identification of a person), the combination of data from several different databases, that in themselves are not personal data, can lead to identification. With regard to big data, it has been asserted that anonymity in data no longer exists (Pentland, 2014). This has severe privacy implications, because when using anonymous data from different databases, an individual can always be identified.
- Even if people manage to not leave any personal data that can be used in data mining, they will be treated as being part of a group with certain characteristics. Moreover, if a group has a certain chance of having a characteristic then all people in this group will be treated as having that characteristic. For example, if adolescents that are converted to Islam in a certain part of town tend to have very orthodox ideas and some of them travel to Syria, then all adolescents converted to Islam in that part of town will be considered vulnerable to radicalisation or even considered radical themselves.

List of references

Abonyi, J. and Feil, B. 2000. *Cluster Analysis for Data Mining and System Identification*. Berlin: Birkhauser.

Fayyad, F., Piatetsky-Shapiro, G. and Smyth, P. 1996. 'From Data Mining to Knowledge Discovery in Databases'. *American Association for Artificial Intelligence*, 17(3):37–54.

Hauge, M. V., Stevenson, M. D., Rossmo, D. K. and Le Comber S. C. 2016. Tagging Banksy: using geographic profiling to investigate a modern art mystery, *Journal of Spatial Science*, 61:1, 185-190.

IARPA (n.d.) *Open Source Indicators (OSI)*. As of 18 February 2018:

<http://www.iarpa.gov/index.php/research-programs/osi>

MacAskill, E., Borger, J., Hopkins, N., Davies, N. and Ball, J., 2013. Mastering the internet: How GCHQ set out to spy on the world wide web. *The Guardian*, 22(2013), p.6.

Oswald, M., Grace, J., Urwin, S. and Barnes G. C. 2018. Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality, *Information & Communications Technology Law*, 27:2, 223-250.

Pentland A. 2014. *Social physics; how good ideas spread; The lessons from a new science*. London: Penguin Press.

Saiya, N. and Scime, A. 2014. 'Explaining religious terrorism: A data-mined analysis'. *Conflict Management and Peace Science*, 32(5):487–512.

Check Appendix: Additional resources for further reading.



CHAPTER ELEVEN

Descriptive statistics

Sarah Grand-Clement, Ben Baruch and Anke van Gorp

See also:



Example(s) of use

The SAFIRE project is a European initiative which has used descriptive statistics to map out the path from radicalisation to violent extremism in groups and individuals. Descriptive statistics were implemented by using data sets concerning culture and attitudes. These include variables such as socio-demographic and economic factors, political and institutional factors, and psychosocial and cultural factors. This has made it possible to understand extremist groups' organisational structure and organise them into clusters (SAFIRE, 2011).

Researchers at Queen Mary University of London and King's College London have carried out a cross-sectional survey to determine the prevalence of vulnerability to radicalisation through factors such as poor mental and physical health, as well as poverty. The survey results were analysed using descriptive statistics and advanced statistical analyses (Bhui et al. 2014).

11.1. Brief description

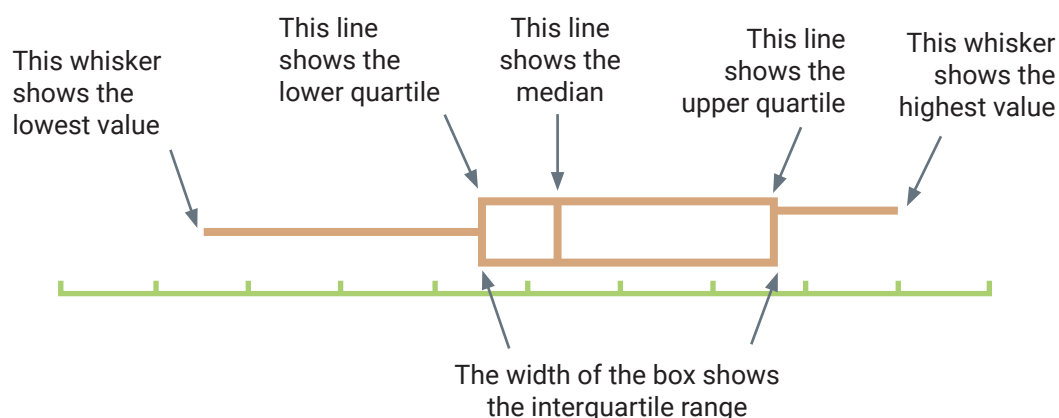
Descriptive statistics refers to a collection of methods used to describe, summarise and present data. Descriptive statistics make it possible to present data in a simplified way that is easily understandable (Bauer, 2009), which is particularly helpful when a data set contains many observations (e.g. mobile phone records or emails sent). This also allows evaluators to look for patterns before using more complex statistical methods. Descriptive statistics

are not, however, used to make predictions or inferences. The following are examples of commonly used descriptive statistics:

- Methods for measuring central tendencies (i.e. averages). These include:
 - Mean: the average value of the data. It is obtained by adding all the values together and dividing by the total number of values. This is a commonly used method to locate the central tendency of a data set.

- Median: the middle value in a data set (e.g. in a data set of 101 observations, the median would be the value of the 51st observation). If there is an equal number of observations, the median is the value midway between the two middle observations (e.g. with 100 observations the median would be the value midway between the 50th and 51st observations). The median is used as a measure of central tendency if the data are skewed (i.e. not a bell-curve). Income is a variable where the median is often used to measure central tendency.
 - Mode: the value occurring most often.
 - Methods for measuring variability (i.e. how spread out or clustered the data points are). These include:
 - Range: this means finding the difference between the highest and lowest value within the data set. This is a useful value to gauge how spread out the data are, but is limited because it only uses two values (the top and bottom).
 - Standard deviation: this is also used in order to understand the extent of variation (spread) of data. However, it is less affected by outliers (values that differ greatly from the majority of the other values). The more spread out the data, the higher the standard deviation value will be (Trochim, 2006). The standard deviation is also used to calculate the standard error, which is used to make inferences from the sample to the population.
 - Quartiles: quartiles divide the dataset into four equal groups. The first 'cut' uses the median to divide the data in half. The median is then found within each half of data. The numbers on which this division falls on are the quartiles. This then shows which values appear in the lowest (25th) percentile, or the highest (75th), for example. If these quartiles are close together, that means that there is not much variability in the data (the dataset is 'compact').
- Results from some of these methods are often best presented using graphs. For example, quartiles can be presented as 'box-and-whisker' plots showing the median, inter-quartile range and outliers. See Figure X for an example:

Figure 1: example box and whisker plot



Source: BBC (2014).

11.2. Purpose

Descriptive statistics are a basic first step in understanding data, typically before undertaking more complex analyses.

Descriptive statistics make it possible to present simple summaries of data sets. For instance, in P/CVE evaluations, descriptive statistics could be used to describe the average or variability of the age of programme targets (if this is believed to be important). They can also be useful to gain an understanding of the context and situation of interventions – for example, by summarising the characteristics of communities deemed to be at ‘high’, or ‘low’ risk of radicalisation in terms average income, education, or more relevant factors such as ‘trust in the police’. Knowledge derived from descriptive statistics can provide basic information to inform the design of interventions through ‘profiling’ individuals, groups or areas.

11.3. When to use it

As stated above, descriptive statistics are a first step used for understanding a data set. Before undertaking more complex analyses, the analyst or evaluator should understand how each variable ‘works’ (i.e. how it is coded), the distribution of a variable and the extent of missing data, and gain a basic idea of how variables relate to one another (e.g. via cross-tabulation). This knowledge helps to inform more complex analysis, for example by understanding whether a variable is skewed and may need transforming, or if missing data is a problem for a particular measure. It is particularly important to understand how potential outcome measures ‘work’ and how much data is missing from them. Basic descriptive statistics can also be used to summarise information arising from other methods, such as focus groups or interviews. Examples of this include collating the

characteristics of focus group members or interview participants in terms of age, gender, ethnicity and so on.

11.3.1. Why?

The use of descriptive statistics has the following advantages:

- They allow for a basic understanding of the data.
- It can help with ‘data cleaning’ – e.g. through identifying mis-classified observations.
- They make it possible to simplify data, particularly if there are many observations in a data set (e.g. producing a simple table showing how many males or females are in a data set).
- They enable a basic understanding of variability in the data (e.g. variation in average levels of education in a sample of neighbourhoods).
- If there are several variables, it is possible to look at basic relationships between variables (e.g. understanding what proportion of females in a data set have previously associated with a radicalised group).

11.3.2. Why not?

There are also disadvantages to the use of descriptive statistics, including the following:

- As their name suggests, descriptive statistics make it possible to describe data. They do not provide any evidence of effectiveness.
- Measures of central tendencies and variability, in particular the mean and range, are particularly sensitive to outlier values. Outliers can skew data and result in an incorrect estimate of central tendency.

- While it is possible to record ordinal data in a numerical form (such as on a Likert scale from 1–10, or very bad to very good) on subjects such as social perceptions or religiosity, it is important to bear in mind that answers to these types of questions are arbitrary.
- If the data were not collected via a random sample, which could be typical of data collected for P/CVE, it is not possible to generalise beyond that data set. For example, finding that 20 per cent of women in a convenience sample had associated with an extremist group does not tell us anything about what percentage of women more generally might have done so. Similarly, if the dataset concerns only one town, this cannot be used to make conclusions about the population of that specific country as a whole. Consequently, this is not a method that can be used to test a hypothesis concerning violent extremism but rather to generate concise information.

11.4. Step-by-step application

The steps required to use descriptive statistics in the context of a P/CVE intervention are detailed below.

1. Determine the variables
 - When focusing on P/CVE, it would be useful to determine what variables or features focus predominantly within the targeted population. This could include factors such as age or income.
 - While numerical data (such as age) can be straightforward to use, it is also possible to use categorical data such as gender, ethnicity and education levels. This data can also be assigned a numerical value. However, some descriptive statistical methods may

not be applicable, even if the variable is numerically coded (e.g. it does not make sense to discuss quartiles of gender, even if gender is coded as 0=male and 1=female in your data).

2. Collect the data
 - If there is no data readily available for use, such as a census, the data will have to be collected. This will require the use of surveys and questionnaires (see the separate chapter on surveys for more information on these).
3. Conduct statistical analysis
 - Input the data into a statistical programme such as Excel, SPSS or Stata.
 - Run statistical tests with the data, ensuring that the tests correspond to appropriate data (numerical or categorical).
 - Generate graphs and charts to summarise the dataset and show the general trends.
 - Review the statistics created and check for any outliers that might skew the results.

11.5. Ethics

If personal data are used in descriptive statistics, evaluators need to adhere to data protection regulations. This means that personal data can only be used for the purpose for which they were collected (purpose binding). Personal data collected for another purpose (for example the implementation of social policy) cannot simply be used for an evaluation unless respondents were informed that the data would be used in that manner.

Numbers and statistics can be regarded as objective and precise and as such might be seen by policymakers or others as highly

informative. However, this depends on the data used and the claims that are made. If the data are not accurate or do not address the precise question required (see the point above about generalisation), statistics only give the illusion of objectivity and precision.

List of references

Bauer, J. 2009. *Statistical Analysis for Decision Makers in Healthcare: Understanding and Evaluating Critical Information in Changing Times*. 2nd ed., Chapter 5.

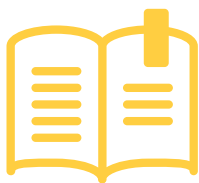
BBC. 2014. Bitesize. Inter-quartile range, cumulative frequency, box and whisker plots. As of 2 September 2017:
<http://www.bbc.co.uk/schools/gcsebitesize/maths/statistics/representingdata3hirev6.shtml>

Bhui, K., Warfa, N., and Jones, E. 2014. Is violent radicalisation associated with poverty, migration, poor self-reported health and common mental disorders?. *PloS one*, 9(3):e90718.

SAFIRE. 2011. *Synthesis report on the results from WP 3 and 4*. As of 2 September 2015:
<http://www.safire-project-results.eu/documents/deliverables/3-4-modeling-culture-indicators-and-intervention-evaluation.pdf>

Trochim, W.M.K. 2006. *Descriptive Statistics*. As of 18 February 2018:
<http://www.socialresearchmethods.net/kb/statdesc.php>

Check Appendix: Additional resources for further reading.



CHAPTER TWELVE

Desk-based research and literature review

Benoit Guerin, Barbara Janta and Anke van Gorp

See also:



Example(s) of use

A study by the Danish Ministry of Foreign Affairs on the prevention of violent extremism in Burkina Faso used desk research, among other methods, to gain an understanding of the potential for extremism and the factors influencing it (Loada and Romaniuk, 2014).

More recently, Davies et al. (2017) carried out a targeted literature review to explore the transferability and applicability of gang evaluation methodologies to counter-violent radicalisation. The study illustrated how exploration of evaluations from established academic fields can be used to inform the development of evaluation practice in relatively novel research areas.

12.1. Brief description

A literature review is a way to obtain an overview of published documents on a given topic, with a view to synthesising, summarising or writing up existing material. Such reviews are an effective way at collecting information at a low cost, saving the time of reinventing an answer or tool that already exists. It is also good research practice – before conducting any primary research, existing evidence should be reviewed in order to shape the approach

to the primary research (e.g. by informing sampling strategies or supporting development of relevant research questions).

Literature reviews are flexible and rely on a range of data sources, including academic literature, online material, discussions with individuals, publicly available statistics, the press, or directories or databases. This can also include so called 'grey literature' such as internal project or programme documentation and reports by different organisations, government institutions,

etc. In the P/CVE context, a literature review could be particularly useful for examining types of programmes that have been implemented, data on programme effectiveness, and lessons learned that may be transferable.

It is important to remember that literature reviews can vary in depth, quality and rigour. In particular, unless the steps in a literature review are clearly stated and are reproducible, one cannot assume that the literature presented offers a complete picture of knowledge on a given topic.

12.2. Purpose

Literature review can be used to find out how much material is available on a specific topic before further research (such as interviews, surveys, etc.) is undertaken. Because it involves reviewing existing material, literature reviews can be useful in determining whether a programme has already been implemented in a given area, or with a certain group. For instance, before developing a new P/CVE intervention, programme developers should run a targeted search examining previous programmes with a similar scope, target group, objectives, etc. This could be helpful in understanding the success factors for a particular intervention, how it was implemented, what it cost and other characteristics. It would also help in avoiding potential intervention pitfalls, as insight into what works and what does not work could be gained from examples of previous programmes, and the extent to which programmes are based on sound ideas of human behaviour.

A literature review can also be useful in assessing 'the state of the art' in terms of academic knowledge on specific aspects of a question. This may include, for example, the most effective ways to engage vulnerable groups in research or the most cost-efficient interventions targeting youth at risk of radicalisation.

It is worth noting that there are different types of literature review, each requiring a slightly different set of skills and resources. Here we outline some of the most commonly used types of literature review.

Literature mapping provides a guide for identifying key concepts, evidence gaps and the type and quality of existing evidence that has been published on a particular topic. In a literature-mapping exercise, the research questions are likely to be broad and the syntheses are typically descriptive. For instance, data mapping could be used to identify the main targets of P/CVE interventions or whether any P/CVE programmes were implemented in schools, without getting into detailed description of these programmes.

Scoping reviews provide a preliminary assessment of the volume and scope of the available literature. Typically, the studies included in a scoping review are not quality assessed and a synthesis of the results would comprise a minimal commentary. A scoping review is typically more structured/defined than a literature mapping. For instance, a scoping review would not only identify P/CVE programmes that were implemented in schools but would also provide basic characteristics of these programmes.

Rapid evidence assessments (REAs) use a systematic approach but the scope of the search and the quality assessment may be restricted. This may be achieved by formally constraining the types of evidence to be sourced, for example on the basis of the type of publication, the language in which the research was conducted, or the time period in which it took place. The synthesis is tabular with narrative commentary. An REA of P/CVE programmes implemented in schools would use a systematic approach to identify all relevant programmes but search and review may be restricted to programmes that were implemented in the last five years

or in secondary schools only. All identified programmes would be systematically reviewed and synthesised.

Systematic reviews are based on extensive search strategy, a comprehensive quality assessment, and if appropriate may include statistical methods (meta-analysis) to analyse and summarise the results of the included studies. For instance, when systematically reviewing P/CVE programmes that were implemented in schools, the first step would be to clearly define inclusion and exclusion criteria to identify and select sources to be reviewed. Selected sources are subsequently reviewed in accordance with an agreed template for data extraction, for example including information on programmes' aims, target groups, implementation methods, funding, etc. The final step is to systematically synthesise the evidence and interpret the findings.

The flexibility associated with literature review means that its results can be expressed or synthesised in different ways, ranging from a short paper to an infographic, a datasheet or a presentation.

12.3. When to use it

Literature can be used before a project (ex-ante) to find out information about the issues that the project or intervention is seeking to address, the main needs of the target group (e.g. radicalised groups or a local community) or the key issues to consider when implementing a programme in a given area.

Literature review can be used as part of an evaluation, especially when it is necessary to obtain a basic answer to an evaluation question within a few days, as this method can yield high returns for a limited amount of time invested (Hague and Wilcock, 2014).

It is important to note that the advantages and disadvantages of this range of methods vary

substantially depending on the type of review used.

12.3.1. Why?

The following are some of the advantages of literature reviews:

- This method is cheaper than others in terms of costs and time since it can provide information within a relatively short timeframe (although this can vary).
- By providing an overview of existing knowledge on a specific question, it can save time by avoiding the replication of past work.
- It can provide new ideas and ways to consider an issue, and can highlight where gaps exist.

12.3.2. Why not?

Disadvantages of conducting literature reviews include the following:

- Literature review can take more time depending on its format (for example, a systematic review is likely to be more time-consuming than a literature-mapping exercise). In addition, initial literature review on P/CVE interventions can yield a large number of research studies and it may be necessary to narrow the search criteria to further define the scope of the topic, for instance by including only interventions conducted in a particular country or targeting particular age groups.
- This method may require advanced research knowledge/expertise, for instance when specific databases are used or systematic review methodology is applied.
- Results from literature review may be biased depending on the scarcity of resources available to do the work, or the number and quality of sources reviewed.

12.4. Step-by-step application

The steps required to conduct a literature review vary depending on the type of review. Typically the standard steps for conducting a literature review include: 1) defining a research question; 2) identifying, selecting and reviewing relevant literature; and 3) synthesising the evidence (production of the output). These steps are outlined below. Note that we set out these steps on the assumption of a reviewer wanting to conduct (at least) a structured search that is reproducible and that answers a specific question.

1. Defining a research question and deciding on the type of review required
 - This step includes defining an appropriate issue to be examined and formulating research questions that will guide a literature review process. To do so, a feasible evaluation question (neither too broad nor too narrow) needs to be defined. For instance, asking, 'What are the successful P/CVE interventions?' may not be sufficiently specific, whereas asking, 'What is the efficiency of P/CVE interventions targeting radicalised youth?' may provide more targeted research findings.
 - Deciding what kind of review is needed (how comprehensive, how systematic it should be) depends on the resources available to do the work, the research question to be answered, planned next steps, and where evidence is likely to be available.
2. Identifying, selecting and reviewing relevant literature
 - In order to identify, select, and review relevant studies, it is necessary to agree on inclusion and exclusion criteria, and to determine search terms. For instance, these may include studies

on P/CVE interventions focusing on youth and exclude P/CVE interventions implemented before 2005. The search terms may include 'school-based intervention', 'P/CVE intervention', 'youth', and combinations of these terms. The search may start broad, with terms such as 'police officers' awareness of radicalisation', and subsequently be narrowed to 'police officers' awareness of the risks of radicalisation among youths from a migrant background'. The next step is to identify sources to be searched. These may include, for example, sources published in English and German in academic peer-reviewed journals, or those made available by implementers of a P/CVE intervention that has been evaluated.

- When the research question has been finalised and the sources have been selected, sources should be assessed for reliability. This includes assessing where study results were published and by whom (e.g. a peer-reviewed academic paper, a blog post or a marketing bulletin by a study sponsor), any bias in the document (for instance, do reports or papers acknowledge other research or points of view?), the type of evaluation results presented (e.g. only positive outcomes or both positive and negative outcomes), and if applicable, the way the research was funded. For instance, it is worth checking whether the author(s) had an incentive to publishing findings in a particular way (e.g. intolerance statistics produced by migrant groups versus official police records on reported incidents of intolerance).
- Sources reviewed should be those that will be most useful in answering the research question(s). It is also an appropriate stage to decide on the

format of the output of the evaluation report (in terms of length, format, etc.). For instance, a desk research report conducted by a police officer would be in a different format and length than a report produced by an agency funding P/CVE programmes.

3. Synthesising the evidence (production of the output)
 - Synthesising findings helps bring together relevant evidence into a cohesive whole, providing a basis from which to make suggestions for future evaluations, research studies and policy formulation or implementation. Collecting findings through this process allows attention to be given to the quality and relevance of the evidence. The synthesis and presentation of results can vary by review type.
 - Important steps to consider at this stage include accurately referencing the sources used (e.g. acknowledging different points of view when sources provide contradictory findings), writing clearly (e.g. avoiding specialised jargon if the report is to be used by a non-specialised audience), and making a coherent argument (Dexter, 2000).
 - Description of findings is followed by a discussion section that aims to synthesise and interpret findings, identify any limitations and explore avenues for further exploratory or analytical work where necessary. The findings are interpreted in light of the quality of evidence and the degree of consistency across studies (i.e. whether all or most of the results are broadly similar) to determine the generalisability of findings. For instance, when all studies report that a specific P/CVE programme is highly effective, it is possible to conclude

that this programme brings positive outcomes to participants and wider society. However, if results regarding the effectiveness of a P/CVE intervention are mixed, with some studies reporting positive results and others reporting no results or negative results, literature reviewers must take particular care when drawing conclusions about the effectiveness of the intervention.

12.5. Ethics

When conducting desk research and literature searches, a number of ethical points should be considered. A general challenge is that analysing documents or secondary material on a programme, policy or intervention may obfuscate some of the complex ethical issues surrounding the study. With regard to the quality of evidence produced, depending on the availability of literature and inclusiveness of the literature search, the results can be skewed, for example when a review focuses on a single research field or professional practice. In addition, plagiarism should be avoided by acknowledging the sources of the information.

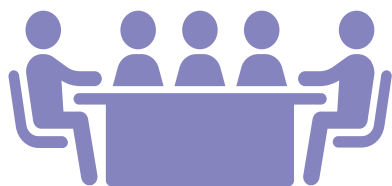
List of references

- Davies, M., Warnes, R., and Hofman, J. 2017. *Exploring the transferability and applicability of gang evaluation methodologies to counter-violent radicalisation*, Santa Monica, Calif.: RAND Corporation, RR-2120-EC, 2017. As of 5 January 2018:
https://www.rand.org/pubs/research_reports/RR2120.html
- Dexter, P. 2000. 'Tips for Scholarly Writing in Nursing'. *Journal of Professional Nursing*, 16(1):6–12. As of 1 May 2014:
<http://uncw.edu/son/documents/dexterphyllischolarlywriting.pdf>

Hague, P. and Wilcock, C. 2014. *How To Get Information For Next To Nothing*. B2B International. As of 18 February 2018: <http://www.b2binternational.com/publications/desk-research/>

Loada, A. and Romaniuk, P. 2014. *Preventing Violent Wxtremism in Burkina Faso*. Ministry of Foreign Affairs of Denmark, Global Center on Cooperative Security, Centre pour la Gouvernance Démocratique Burkina Faso. As of 18 February 2018: <http://www.globalcenter.org/wp-content/uploads/2014/07/BF-Assessment-Eng-with-logos-low-res.pdf>

Check Appendix: Additional resources for further reading.



CHAPTER THIRTEEN

Focus groups

Kate Cox and Anke van Gorp

See also:



Example(s) of use

The UK Institute for Strategic Dialogue used focus groups in its research project titled 'Stepping Out: Exit Strategies from Extremism' (Choudhury, 2009). The study aimed to identify practical ways to support de-radicalisation processes in Britain, Germany and Denmark. This involved engaging former radicals who had rejected extremist ideologies and organisations. Focus groups were used to examine reasons for joining, triggers for leaving and the barriers individuals faced in trying to disengage with extremist organisations, while also considering the role and contribution that former extremists themselves might play in a network that could support the process of leaving these organisations.

After the 7/7 bombings in London in 2005, homegrown radicalisation was widely recognised as an issue in the UK, with Muslims born and brought up in the country being responsible for the attacks. In order to fill knowledge gaps on the issue of radicalisation within the Muslim community in the UK, a group of scientists from London universities ran a project that involved focus group consultations with Muslim community members (Ghosh et al. 2013). Two focus groups with six men and ten women, aged between 22 and 56, were established to determine the meaning of radicalisation to Muslims, gather experiences of the impact of the concept of radicalisation on the wellbeing of the Muslim community, understand more about the sociological and psychological processes that lead to radicalisation, and gather in-group perspectives on how to tackle radicalisation as a means to promote social cohesion.

13.1. Brief description

A focus group is a form of qualitative research in which participants (e.g. local community

representatives or police officers engaged in P/CVE programme delivery) are asked about their attitudes towards a programme or concept, for instance the effectiveness

of a P/CVE programme or the identification of a new intervention. Focus groups usually consist of 6 to 10 participants and are guided by a moderator (usually a researcher) who facilitates discussion. Participants are usually individuals with a similar background, professional affiliation or responsibilities (e.g. P/CVE programme directors from across different regions). Data is shaped and refined through group interaction.

There are many different types of focus group, including:

- **In-person focus groups:** conducted face-to-face with participants and the moderator gathered around a table. Discussions may be recorded by audio or videotape or observed unobtrusively for note-taking purposes.
- **Online focus groups:** conducted over the internet with participants posting their answers and asking questions on the online discussion platform. This can be helpful when participants are geographically dispersed (e.g. chief constables from different regions).
- **Teleconference focus groups:** conducted over the phone with participants dialling-in/ connecting at a specific time. Similar to an online focus group, participation in the teleconference focus group allows participants to connect from their location, making it a suitable method when consulting a range of stakeholders from different locations.
- **Two-way focus group:** one focus group observes another focus group and discusses their interactions. For instance, headmasters observe a discussion between P/CVE programme developers on how the P/CVE programme should be implemented in schools. Subsequently, during the second focus group, headmasters provide feedback on the

implementation planning and provide suggestions on how to improve it.

- **Duelling moderator focus group:** two moderators deliberately take opposite sides on the issue under discussion. For instance, during a focus group discussion on appropriate punitive measures for a violent behaviour, one moderator could represent a very strict approach, suggesting imprisonment for any form of violent behaviour, whereas the other moderator could express the view that only the most severe cases should be imprisoned and other offenders should be given community service.
- **Mini-focus groups:** groups consist of 4 to 5 members rather than 6 to 12. A smaller number of focus group participants can be helpful when consulting individuals that may be difficult to control or recruit, or if they are few in number. Another advantage of a mini focus group is that it increases each participant's opportunity to share insights and views. On the other hand, a smaller group limits the total range of experiences.

13.2. Purpose

As with in-depth interviews, focus groups are useful for research that aims to explore, pilot or refine a programme concept; identify participants' goals, expectations and views of the efficacy of an existing or proposed programme; document experiences of implementing a programme; or describe differing outcomes across participant groups. In the P/CVE context, focus groups can be useful to discuss ideas about a new P/CVE programme, to obtain practitioners' and the local community's views on the perceived facilitators and barriers involved, or to assess a programme's effectiveness and efficiency across a range of stakeholders.

13.3. When to use it

Focus groups are primarily used for data collection. Following the focus group discussion, participants' observations tend to be analysed (e.g. for similarities and differences between views of particular participants representing different types of stakeholders, such as police, teachers, youth, etc.) and used as a basis for report writing (e.g. formulating conclusions and recommendations for further action). Focus group findings are also sometimes used as a developmental framework for additional data collection such as surveys. For instance, a particular P/CVE issue could be discussed during a focus group to formulate ideas for survey questions or to test a survey instrument.

13.3.1. Why?

Focus groups have a number of advantages as a research method, including the following:

- They are more time efficient than other research methods (e.g. interviews) as views of a range of participants are investigated during the same discussion.
- Focus groups are less expensive than other research methods (e.g. interviews) since relevant stakeholders are consulted during the same session.
- Group dynamics often draw out information that may not have been anticipated by the researcher and may not have emerged from individual interviews. For instance, when focus group participants represent a diverse range of stakeholders, the views of one participant (e.g. a representative of a youth organisation) may be contested by another participant (e.g. a local community representative), and this may lead to a lively discussion which takes a different direction than initially planned by the focus group's moderators.

- A discussion-based format has the potential to generate a wide range of opinions and promote the exchange of ideas and information, as participants have a chance to discuss issues with others who may have a different view point on the same subject (e.g. youth leaders can learn from police officers and vice versa).

13.3.2. Why not?

There are also disadvantages to the use of focus groups. These include:

- Participation bias: the selection of participants often draws on purposive sampling and includes a limited number of participants, meaning that findings are not generalisable to a wider population.
- Moderator bias: there is a risk that the moderator intentionally or inadvertently injects their personal biases into the discussion, leading to inaccurate results. For example, if the moderation is conducted by the local authority's research team, the moderator could be seen as a representative of the authority rather than an objective researcher.
- Social desirability bias: there is a tendency for respondents to answer questions in a manner that will be viewed favourably by others. For instance, if teachers and students participate in the same focus group discussion, students may answer questions in a way that they see as desirable to teachers.
- Groupthink: there is a tendency for participants to try to minimise conflict and reach a consensus decision without critical evaluation of alternative viewpoints and by actively suppressing dissenting viewpoints. For instance, if participants are of different ages (e.g. young and older community leaders), the older participants may have

more power to suppress the views of younger participants.

- Lack of anonymity: participants may be more reluctant to disclose sensitive information in a group context. For instance, prisoners may be reluctant to discuss their offences in a group setting.
- Lack of depth: focus groups do not allow for as much detail as individual interviews in their analysis of issues under discussion. Participants may only be able to highlight an issue in broad terms without elaborating on it through deeper discussion. For example, certain aspects of radicalisation may be mentioned without discussing in depth how they affect particular societal groups.

13.4. Step-by-step application

1. Clarify expectations

- Determine the purpose and scope of the focus group. For instance, at the beginning of discussion, briefly explain the purpose of the meeting and rules of engagement.
- Determine the range of participants needed and how they should be subdivided into different focus groups.

2. Select and invite participants

- Choose a homogenous selection of participants in terms of job classification and level of education. For example, when discussing de-radicalisation programme have separate focus groups for young offenders and for community workers. Note that the homogeneity of the group will limit the scope for interaction of different perspectives during the session. On the other hand, another option is to combine the perspectives of community workers and police or community leaders in one focus group.

- Do not select participants who are part of an existing group (e.g. colleagues or friends). For instance, do not invite several teachers from the same school where a P/CVE programme was implemented; instead aim to invite teachers representing different schools.
- Send email invitations and, if necessary, follow up with telephone calls.

3. Appoint a moderator

- Select a moderator who is not associated with the programme under discussion (external is preferable) and is not a representative of a particular interest group (e.g. a P/CVE programme developer).

4. Create a protocol

- Determine a set of questions before the focus group is held.
- Write a semi-structured protocol, which should rely on open-ended questioning and progress from the general to the specific. For instance, ask who were the main beneficiaries of a P/CVE intervention and follow up by asking what the benefits were for each group or individual.
- Questions should fall into five general categories: opening questions, introductory questions, key questions (focusing on the main areas of concern), and concluding questions. Plan specific timing for asking these questions (e.g. 3 minutes for opening questions, 8 minutes for introductory questions, etc.).
- If relevant, prepare introductory material to be shared with the participants to shape the discussion. This material needs to be understandable for the audience and avoid introducing bias to the discussion.

5. Open the focus group

- Welcome participants, explaining the ground rules and purpose of the focus group. If the participants do not know each other, it may be helpful if they briefly introduce themselves.
 - A warm-up exercise or ice breaker usually helps facilitate engagement with the participants.
6. Moderate the focus group
- Respond to, probe and follow up with participants to gain a clear understanding of their responses. For instance, when participants discuss a specific aspect of a P/CVE intervention, probe for more detail that is relevant to the objectives of the discussion.
 - Ensure that one person does not dominate the discussion (e.g. the person with the highest level of seniority, such as a chief constable or head teacher).
- 6.1) Record the discussion
- Take notes and audio recordings of the discussion.
7. Analyse findings
- After the workshop, analyse the notes or transcripts to draw conclusions about the programme discussed. This could be thematically, by question asked, or in relation to the main study research questions.
 - Identify points of agreement and dissenting views across groups and individuals in the post-focus group report (e.g. community leaders and teachers having similar views on a P/CVE programme but differing from police officers).
 - Use findings as a developmental framework for additional data collection such as surveys (this step is optional).

13.5. Ethics

Anonymity is perhaps the most ethically contentious aspect of focus groups. The following steps may be taken to address this issue:

- Provide participants with full information about the purpose and use(s) of their contributions in advance of the focus group discussion, e.g. by explaining that their contribution will only be used to further develop a P/CVE programme.
- Ask participants to read and sign a consent form prior to the focus group.
- At the start of the focus group, clarify that each participant's contributions will be shared with the others in the group as well as with the moderator. In case of a two-way focus group, inform participants in the first focus group that views will also be shared with participants in the second focus group.
- Encourage participants to keep confidential what they hear during the meeting.
- Anonymise data from the group during the research and report-writing processes. For instance, when describing participants' views, refer only to their generic job titles (e.g. instead of saying 'science teacher from secondary school', say only 'teacher').

List of references

- Choudhury, T. 2009. *Stepping Out: Supporting Exit Strategies from Violence and Extremism*. London: Institute for Strategic Dialogue.
- Ghosh, P., Warfa, N., McGilloway, A., Imran, A., Jones, E. and Bhui, K. 2013. 'Violent Radicalisation and Recruitment to Terrorism: Perspectives of Wellbeing and Social Cohesion of Citizens of Muslim Heritage'. *Sociology Mind*. 3(4)290–297.
- Check Appendix: Additional resources for further reading.



CHAPTER FOURTEEN

Interviews

Barbara Janta and Anke van Gorp

See also:



Example(s) of use

In 2008, the Change Institute conducted a study on violent radicalisation for the European Commission. It used primary fieldwork interviews in four countries to explore the ideologies, beliefs and individual narratives leading to violent radicalisation (Change Institute, 2008).

An Austrian programme, Mothers Schools, Sisters Against Violent Extremism (SAVE), has used interviews to understand the process of radicalisation that young people undergo. Subsequent elements of the programme focused on teaching family members (mothers) how to recognise and react to potential early warning signs in their children's behaviour.

The aim of the UK's programme Tackling Radicalisation in Dispersed Societies (TaRDIS) was to understand and tackle radicalisation in dispersed societies. In order to achieve this aim, the programme applied the Intelligence through Neighbourhood Security Interviews, conducted with members of the local community (Lewis, 2013). This allowed the programme team to develop a deeper understanding of the potential for dissatisfaction among local residents and the potential for those individuals to be radicalised towards violent extremist viewpoints.

An academic study by Gambetta and Hertog (2009) used interviews to examine the link between an engineering education and propensity to become an Islamic radical.

14.1. Brief description

Interviews are a common technique used to collect data when conducting evaluations of P/CVE interventions (see examples in the section on resources). They consist of an

in-depth conversation to gather information which relates to the interviewee, which can be more or less structured depending on preference. This may include their position in a radical group or an organisation that works to prevent or counter violent extremism), or

their views, thoughts or opinions on a certain topic (e.g. which elements of the P/CVE intervention in which they participated were more or less helpful and why). The average length of an interview will vary depending on the group or individual being interviewed. It would not be realistic to expect to spend two hours interviewing a high-level public official responsible for a nation-wide P/CVE programme. In this case, an interview may be limited to 30–40 minutes. On the other hand, a conversation with an implementing officer may be more in-depth and take more time.

Interviews can be undertaken with different degrees of flexibility. They can be:

- Structured: following a specific questionnaire with all interviewees being asked exactly the same questions.
- Semi-structured: combining specific questions with some freedom to ask unplanned follow-up questions.
- Unstructured: consisting of a free-flowing conversation on a given topic.

14.2. Purpose

The data obtained through interviews is qualitative in nature and enables the person leading the interview to gain a clearer understanding of the opinion, beliefs, experience or thoughts of the person interviewed (Bryman, 2001, p.319). These often focus on a project, programme or institution that the interviewee is either affected by (e.g. as a recipient of an integration or de-radicalisation programme) or invested in (e.g. as an implementation officer in a charity helping young people at risk of radicalisation).

Interviews can also be used to explore change over time in 'the reasoning, conceptions and representations of the persons questioned on a project or programme', or individual practices in

using a service or engaging with a programme (European Commission, 2013, p.121).

14.3. When to use it

While interviews can be used at all stages of an evaluation, it is particularly useful to conduct them early on. At the stage when the research is being designed, interviews may enable the research team to get acquainted with a subject, for instance by identifying relevant agencies or actors involved in implementing P/CVE programmes at a local level and establishing links and relationships between them. They may also enable the research team to develop research tools (e.g. testing the relevance of survey questions). At the data collection stage of research, interviews can enable the team to identify issues to be explored with other methods, or can provide insights into previously identified aspects (Ling and Villalba van Dijk, 2009).

14.3.1. Why?

The following are some of the advantages of interviews:

- Interviews help to obtain rich, detailed information in a relatively inexpensive manner. It can also be used in formative evaluations to test the logic of an intervention (for instance a new P/CVE programme), or to help fine-tune existing programmes or interventions (e.g. when implementing a programme in a different region or country).
- Interviews are especially useful to obtain:
 - Descriptions, perceptions or perspectives on a programme or institution (e.g. perceptions of P/CVE programme participants and their families), notably to complement information gathered in other ways (e.g. statistical data);

- In-depth, candid, information about the perceptions of diverse individuals (e.g. police officers' perceptions of particular disengagement programmes);
- In-depth information on how the outcomes or implementation of a specific programme differ across sites or groups (e.g. comparing views of youth leaders from various locations participating in the same type of P/CVE programmes);
- Interviews can be used when recommendations on specific parts of a programme or institution need to be formulated (e.g. consulting on programme options with potential participants, programme leaders and representatives of the wider community).
- They complement other research approaches or methods, for instance to help develop surveys or focus groups once it becomes clear what the main issues are in a given area. They can also enable in-depth exploration of new ideas and insights (e.g. interviewers can follow up on specific aspects of a P/CVE programme to explore it in depth).
- Interviews make it possible to engage in relationship-building with major actors or stakeholder organisations (e.g. community leaders, radicalised individuals). They provide an opportunity to clarify questions or terminology when required.
- They can be used to identify individual concerns which may not be raised in focus groups due to participants being afraid of discussing sensitive P/CVE issues in a group setting.
- Interviews enable researchers to quickly obtain insights into the practices and norms of a given community (e.g. seniority levels in radical groups).

14.3.2. Why not?

Some disadvantages of interviews as a research method include the following:

- It is difficult to ensure that the findings are unbiased or are valid on a larger scale, and a large number of respondents is required in order to limit bias. For instance, to have a good understanding of a programme delivered to radicalised individuals, it would be necessary to speak to programme participants and those who delivered/implemented the programme, as well as family and wider community representatives across different locations where the programme was implemented.
- It can be difficult to select respondents when a range of viewpoints need to be represented. Selecting the 'right' respondents and making contact with them can be time-consuming, in particular if potential interviewees are from hard-to-reach populations. Typically, an initial long list of potential respondents related to a P/CVE programme would be drafted. A final list of interviewees would then be made based on their availability.
- Interviews can be time-consuming in terms of travelling (if conducted in person), conducting the interview and synthesising insights from a range of stakeholders.
- The interviewer–interviewee relationship can be undermined if the mode of interaction (i.e. online, by phone) is not well received. For instance, conducting interviews by phone would be acceptable when interacting with professionals, such as programme developers, police officers, policymakers, etc. However, in-person interviews would be more appropriate when collecting data from service users, such as radicalised individuals or their family members.

- This method requires skilled interviewers to ensure that the interviewee feels comfortable sharing information about their programme or project. In order to gain access and build rapport with interviewees such as members of violent groups, interviewers may need prior buy-in and agreement from senior people in a group. Interviewers must develop mutual trust (Becker and Bryman, 2004) and have 'cultural competence' in order to get access to potential interviewees (e.g. a male white interviewer may not be able to gain access to and conduct an interview with a radicalised woman from an ethnic minority).

14.4. Step-by-step application

The immediate next steps required to use this method to assess the effectiveness of a project in the area of P/CVE are detailed below.

1. Review relevant literature or documents related to P/CVE
 - Select a shortlist of topics to be covered by the interview (e.g. who are the main beneficiaries of the programme, how the programme operates, how it is funded, what has already been written on the subject, etc.).
 - Identify specific potential interviewees including funders, relevant stakeholders (e.g. police, community/civil society groups) and academics, as well as potential contacts within organisations or groups of people.
2. Select interviewees
 - Identify the groups that potential interviewees belong to (e.g. beneficiaries of the programme, agents implementing the programme, etc.). The sample can be based on characteristics of the target population (e.g. age or gender of programme beneficiaries), or of the programme (e.g. location where it is implemented).
3. Prepare the interview questionnaire
 - Determine how many interviews need to be/can be conducted. Note that the number of interviews may need to be revised if saturation is reached, which is when interviewees are no longer providing new information.
 - One way to select interviewees can be to look at lists of attendees at relevant events, mailing lists for relevant organisations, etc. Please note that it is important to avoid potential bias when selecting interviewees (e.g. only the closest geographically or easiest to access), and to ensure a representative sample. Representative here means covering all stakeholders, rather than in the statistical sense. For example, when assessing the benefits/drawbacks of a programme, it is best to gauge views of a range of stakeholders (e.g. programme participants, wider community, police officers, policy-makers etc.).
 - Plan the format of the interview: semi-structured, structured or unstructured (see Becker and Bryman, 2004).
 - Select the venue or mode of communication: in person, by phone, etc.
 - Write the interview questionnaire and estimate the time needed to cover all questions. In doing so, be sure to:
 - Introduce yourself and what you do, and explain to the interviewee why you need their cooperation, how the information will be recorded and analysed, how the anonymity of interviewees will be protected (if relevant), etc.
 - Have a number of key questions which are relevant to your interviewee (plan

the number of questions depending on how much time is available). Some questions could be the same across a range of interviewees (e.g. views on the impact of a programme on community engagement). However, some aspects may be discussed only with some interviewees. For example, skills needed for successful programme may be delivery discussed only with programme implementers, while financial constraints faced by community organisations may be discussed only with leaders or members of these organisations.

- You can also use 'probing' or 'prompting' questions to encourage your interviewee to reflect on their practice, ask for specific details to follow up on their answer, etc. Prompts can include 'Could you tell me more about ...?' or questions such as: 'You said that the main success factors for delivering this P/CVE initiative are X, Y and Z. Are there any other important aspects that have an impact on how you deliver this programme?'
- Do not forget to include a 'closing' question asking them whether they have any additional comments. For instance, conclude by saying that you have explored all the questions that you prepared in advance, and ask the interviewee whether they have any other comments or observations, or if they want to raise any issues that have not been discussed.
- Make a note to thank your interviewee for their time.
- After selecting interviewees and planning the format, contact them and schedule a suitable time to talk. Send out invitations to interviewees. Please note that owing

to constraints in their diaries, it may be necessary to need to be flexible in terms of time, venue, etc. For instance, interviewers might be asked to come to the offices of interviewees (e.g. for senior policymakers) or conduct interviews over the weekend (e.g. with voluntary community leaders).

- Select an appropriately skilled interviewer to conduct the interview.

4. Conduct the interview

- Be sure to have a solid understanding of the subject matter. If in doubt on any aspects related to the issue being examined, do not be afraid to ask questions, or ask for clarification.
- The relationship with the interviewee is an important aspect to manage. It is crucial not to influence the interviewee or to nod or frown at their responses. For instance, when interviewing religious leaders involved in delivery of P/CVE programmes, interviewers should not expose their own agreement or disagreement with the interviewee's religious views. Taking notes and recording the interview can be useful, but may create tension and limit the candidness of the interviewee. Consent for recording the interview/taking notes should always be sought before starting.

5. After the interview

- Produce a transcript of the interview, if necessary. This can be a longer narrative-style summary or a bulleted list of key points.
- Analyse and synthesise findings from all the interviews conducted throughout the project.
- Share the transcript or notes with the interviewee for validation and ask

permission to use any quotes – see section on ethics below.

14.5. Ethics

When conducting interviews, the following ethical procedures should be adhered to:

- Inform respondents about the goal of the project and who has access to their data.
- Anonymise the transcripts and store interviewees' contact details separately from transcripts. For instance, all contact details for interviewees can be kept in a password-protected Excel file. This file will also include numbers/codes assigned to each interviewee. Interview notes would be stored as separate word files identifiable only by numbers/codes.
- Sending a transcript to the interviewee and asking him/her to check it and give permission to use it for the research ensures that the interviewee has given his/her consent. Typically, when conducting interviews with senior officials, the interviewer may be asked to send them a draft version of the interview transcript for review. Interviewees who review a transcript often give valuable extra information and can correct misunderstandings (bear in mind that allowing review of transcripts may not be feasible in practice).
- When conducting interviews with vulnerable groups, it may be necessary to develop information material on the background to the project, as well as a consent form to gain informed consent from participants. For instance, when arranging interviews, it may be useful to send interviewees a one-page document summarising the study aims and explaining the purpose of the interviews.
- Harm to participants, whether emotional or physical, has to be avoided. Interviewers have to use their intuition and experience to determine whether or not to end or interrupt an interview. For instance, when discussing the process of victimisation, interviewers may need to move to another question, or finish the interview all together, if the interviewee becomes distressed. The amount of time requested of respondents and the emotional burden should not be excessive.
- When conducting interviews with vulnerable groups, the issue of power has to be considered. The relationship between the participant and the interviewer should not influence the answers of the participant. For example, interviewees may feel more comfortable being interviewed in familiar settings (e.g. in a community building or school, rather than coming to the interviewer's office).
- When conducting interviews there may also be a risk to interviewers if challenging issues are explored. This may impact on mental health and wellbeing of research staff over time, particularly where they may personally relate to some of the issues raised. This is particularly pertinent for interviews where a personal connection is likely to develop between interviewer and interviewee.

List of references

- Becker, S. and Bryman, A. 2004. *Understanding research for social policy and practice. Themes, methods and approaches*. The Policy Press and the Social Policy Association.
- Bryman, A. 2001. *Social Research Methods*. Oxford: Oxford University Press.
- Change Institute. 2008. *Studies into violent radicalisation: The beliefs, ideologies and*

narratives. As of December 2014: http://www.changeinstitute.co.uk/images/publications/changeinstitute_beliefsideologiesnarratives.pdf

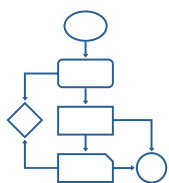
European Commission, Directorate-General for Regional and Urban Policies. 2013. *Evalsed Sourcebook: Methods and Techniques*. European Commission.

Gambetta, D. and Hertog, S. 2009. "Why are there so many Engineers among Islamic Radicals?". *Archives Européennes de Sociologie*, L(2):201–230.

Lewis, C., 2013. 'The Safer Sutton Partnership Service: Blueprint for the Future?'. *The police journal*, 86(1), pp.15-28.

Ling, T. and Villalba van Dijk, L. 2009. *Performance Audit Handbook*. Santa Monica, CA: RAND Corporation.

Check Appendix: Additional resources for further reading.



CHAPTER FIFTEEN

Logic models/ theory of change

Joanna Hofman and Anke van Gorp

See also:



Example(s) of use

Prevent forms part of the UK Government's counterterrorism strategy and aims to tackle the causes of violent extremism at both individual and community level. Underpinned by significant financing, it has been crucial to understand what outcomes have been achieved by the local authority-led activities funded under the Prevent programme. A methodological paper was produced by the Tavistock Institute of Human Relations to inform and guide the national evaluation of local initiatives. The paper includes a draft logic model with some explanatory guidance on which elements should be included in the evaluation (Junge et al. 2009).

In 2011, the Tavistock Institute of Human Relations carried out an evaluation of the London Borough of Tower Hamlets' Prevent projects. As the emphasis of the evaluation was on learning, the evaluators used a theory of change approach in order to gain an understanding of the successes and challenges of the intervention. The theory of change approach helped to identify the pathways to change and which factors may enable the intervention to make a difference (Iacopini et al. 2011).

15.1. Brief description

Logic model is a generic term that describes various representations of projects, programmes or other interventions that link their key components (e.g. inputs, activities,

outputs, outcomes and impacts) to intended objectives. These elements are related to each other by assumptions of causality ('if..., then...'). These models can be relatively simple (such as *intervention logics*) or more complex (such as the *logical framework* or *theory of*

change models). Logic models and ToC models are examples of theory-based approaches to evaluation, alongside other theory-based approaches such as contribution analysis, realist evaluation and policy scientific approaches.

15.2. Purpose

Logic models provide a framework for developing an implementation plan, as well as designing and conducting an evaluation. There are significant differences in how and for what purpose various models can be used in an evaluation. These are explored below.

Intervention logics are relatively simple models that graphically illustrate programme components, or the steps in a process. The intervention logic takes a narrow but descriptive look at the relationship between inputs and results and summarises a complex theory into basic categories. In the case of a P/CVE intervention, the intervention logic may also include the rationale or the needs that the intervention is trying to address and follows the process through to the expected outcomes.

Logical frameworks (or log-frames) present the intervention logic in table format and add information on how the achievement of objectives can be demonstrated through indicators, how these can be obtained (sources of verification), and what assumptions and risks were identified.

Theory of change models link a policy intervention's context, activities and results in order to explain how and why the desired change is expected to happen. A ToC explains (rather than simply describing, as in other models) the causal relationships between context-input-output-outcome-impact in order to understand the combination of factors that will ultimately lead to the expected impacts. As such, a ToC takes a wider view of a desired change by considering a contextual situation, assumptions (or pre-conditions that need to be

met to allow the change) and related risks, as well as intended and unintended effects.

15.3. When to use it

Logic models can be used at the outset to plan an intervention. This may help establish the necessary inputs in order to achieve the desired outcomes of a P/CVE intervention from the outset. When conducting an evaluation, logic modelling is usually undertaken in the initial phase of the research. If the programme theory has not been clearly articulated by the intervention/developer, it is reconstructed in order to identify hypotheses to be tested in the course of data collection and analysis.

15.3.1. Why?

There are a number of advantages to using an intervention logic, including the following:

- It provides a helpful visual 'snapshot' of a project/programme
- It can be flexible and adaptable in its use
- It can be a useful tool to communicate the purpose of the project/programme, the expected results and the actions expected to lead to the desired results
- It can become a reference point for everyone involved in the project/programme
- It improves expertise in planning, implementation, and evaluation
- It involves stakeholders, enhancing the likelihood of resource commitment
- It identifies potential obstacles to a programme's operation so that these can be addressed early on.

Using a logical framework also has various advantages, including:

- It is simple, visually accessible and easy to understand

- It provides a graphical representation of the results chain
- It encourages the examination of risks and assumptions
- It requires analysis of whether the objectives are measurable
- It links problem analysis to objective setting
- It can be applied in a participatory way.

A ToC approach has the following advantages:

- It takes into account uncertainty and diversity of relations
- It helps handle complexity without falling into oversimplification
- It helps to understand and explain the mechanisms of an intervention, establish causality and tell a credible story
- It makes assumptions explicit and analyses them critically
- It can be dynamic (more open to modifications)
- It includes stakeholders in the process of defining/amending ToC and establishes common principles and vocabulary
- It designs more realistic plans of action
- It creates more meaningful evaluations.

15.3.2. Why not?

Using an intervention logic has the following disadvantages:

- It does not represent the full complexity of reality, but simplifies in order to aid understanding. In doing so, it ignores a variety of external factors that influence success or failure in a P/CVE intervention
- It can be difficult to formulate due to the involvement of various decision makers with differing interests.
- It can be applied mechanistically, limiting its ability to reflect the evolution of the project.

There are also a number of drawbacks to using a logical framework:

- It is simplistic, static and inflexible
- It does not take into consideration complex interdependencies between different factors
- It over-specifies/oversimplifies the objectives, leading to rigidity
- It might overemphasise quantifiable data
- It may overlook unintended effects
- It is more difficult to apply to complex projects and programmes with multiple aims.

Finally, the disadvantages of a ToC approach include the following:

- It takes time and effort to develop and reach consensus among various stakeholders
- The result (visualisation) might not be entirely clear for those who were not involved in the process of developing it.

15.4. Step-by-step application

Ideally, a logic model will already have been developed at the programme design stage, but sometimes it is necessary to reconstruct or refine it after the programme has been established. In developing the logic model it is important to identify key components of the programme and relationships between them, and validate them with key stakeholders. It should be noted that developing a logic model and ToC in particular is a highly iterative process that requires consultations with key stakeholders.

Intervention logic steps

1. Problem statement: articulate the problem that programme is trying to solve or the needs that it is trying to address (e.g. increase resilience).
2. Goals: define what the programme is trying to accomplish.

3. Resources and activities: identify available resources and how these will be used to implement the programme and achieve its goals (activities).
 4. Outputs: identify measurable, tangible and direct products of programme activities (e.g. recidivism or exit rate).
 5. Outcomes: determine how outputs make a lasting change and contribute to programme goals (e.g. long-term recidivism or exit rate).
5. Define indicators:
 - The indicators should be objectively verifiable.
 - The indicators should be defined in relation to the goal/purpose/ (component) objectives identified.
 - The indicators should correspond to all the criteria likely to show that what was expected was also produced, thus providing tangible proof of success.

Logical framework steps

1. Define the goal/purpose/objectives:
 - The global objectives or goals serve as a basis for assessing an intervention in relation to longer-term and more diffuse effects (or global impacts).
 - The purpose or component objectives can be defined in terms of intermediate, specific and/or operational objectives.
2. Describe the inputs, activities and outputs:
 - Inputs: the resources put towards the programme/project
 - Activities: the actual tasks to be performed
 - Outputs: the product or service delivered.
3. Define assumptions and risks:
 - Risks and assumptions should be identified in relation to the stated purpose/(component) objectives, activities and outputs.
4. Define sources:
 - The source or sources of verification (SoVs) should be defined in relation to each indicator.
 - The SoVs are described by detailing the means and sources of information needed to obtain the required data.

Theory of change steps

The figure overleaf outlines the steps involved in a ToC approach.

15.5. Ethics

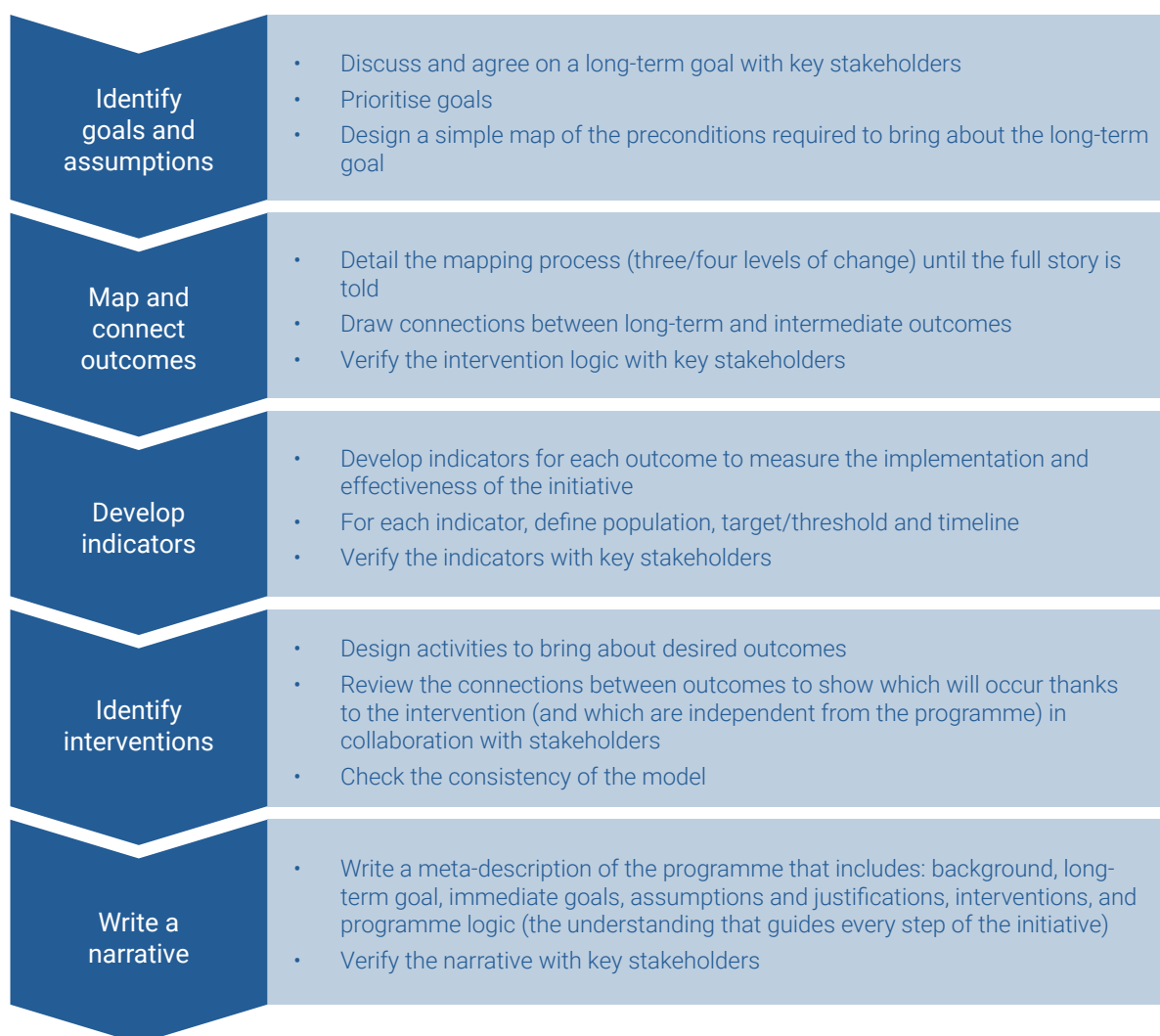
When creating logic models in evaluation the following ethical points should be considered:

- Ethical issues may arise depending on the methods used to gather data. See for example the chapters of this document on interviews, surveys, focus groups, literature review and case studies.
- An ethical issue specific for all theory-based evaluation might be that it is difficult to reconstruct the programme theory without falling prey to hindsight bias (Roese and Vohs, 2012). Stakeholders might be affected by what they have experienced and learned during the intervention and therefore may not be able to correctly reconstruct the assumptions and ideas they had during the design of the intervention under evaluation. Moreover, if they have invested in a project already, the 'sunk cost fallacy' might also cloud stakeholders' judgment (Dobelli, 2013).

List of references

Dobelli, R. 2013. *The art of thinking clearly*. London: Sceptre.

Figure 2: Steps in theory of change development



Iacopini, G., Stock, L. and Junge, K. 2011. *Evaluation of Tower Hamlets Prevent Projects Final report prepared for the London Borough of Tower Hamlets*. The Tavistock Institute. As of 18 February 2018: http://www.tavinstitute.org/wp-content/uploads/2012/12/Tavistock_Projects_Tower-Hamlets-PVE_evaluation_final_report.pdf

Junge, K., Hills, D., Cullen, J. and Das, J. 2009. *Report: A Methodological Paper to Inform the Future Evaluation of CLG-funded Local Authority Preventing Violent Extremism Work*. The

Tavistock Institute. As of 18 February 2018: http://www.tavinstitute.org/wp-content/uploads/2013/01/Tavistock_Report_Methodological-paper-PVE-work.pdf

Roesse, N.J. and Vohs, K.D., 2012. Hindsight bias. *Perspectives on psychological science*, 7(5), pp.411-426.

Check Appendix: Additional resources for further reading.



See also:



CHAPTER SIXTEEN

Longitudinal studies and data analysis

Catherine Saunders and Anke van Gorp

Example(s) of use

A study examined the influence of collective strain on support for violent extremism among adolescents (Nivette, 2017). Authors used data from two waves of an ongoing prospective longitudinal study on the development of aggressive and other problem behavior based on a cohort of children who entered 1 of the 56 primary schools in Zurich in 2004. The study found that adolescents who already advocated for violence and rule breaking were more susceptible to extremist violent pathways.

16.1. Brief description

The defining feature of a longitudinal study is that it collects or uses information or data from more than one point in time. However, within this broad definition there is a spectrum of possible approaches. Longitudinal approaches are in contrast with cross-sectional studies, which provide a snapshot at a single moment. Details of some longitudinal research designs are given below:

- **Cohort study:** in a cohort study, subjects are selected on the basis of their belonging to a specific group of individuals or their exposure to something (e.g. whether they were born in a specific year or live in a particular area), and are followed up over time. The purpose might be to study development more generally, or to understand whether, having been exposed to something, a given outcome develops.
- **Intervention study:** in an intervention study, subjects are selected and then

allocated (often by investigators) to groups in which they receive or do not receive the intervention. This allocation may be at random (a randomised control trial), or may use another systematic approach, or may be based on some other characteristic (e.g. whether eligible for a particular welfare payment or not).

- **Repeated cross-sectional study:** cross-sectional studies examine the same population over time. The same people are usually not included in each wave of the study, but the same data collection approach can be taken at multiple time points with the same group. One example might be collecting data on all Year 10 pupils in one city, for several consecutive years meaning that each time it is the current Year 10 pupils, to understand changes over time between cohorts.
- **Panel study:** in a panel study, repeated data collection is carried out over time, but the 'panel' nature of the design means that the same individuals (or units) are measured at different time points of the study. It is often the case that some people cannot be followed over time, and so new individuals may be recruited to replace them. This design is different from a cohort or intervention study as subjects are not selected on the basis of their exposure, nor are they assigned to an intervention.
- **Ecological studies:** ecological studies look at grouped data from populations, rather than data from individuals. An example might be research on the attitudes of residents to community cohesion in a particular city which, instead of using information from separate individuals living in the city, only uses summary data on attitudes collected in different neighbourhoods and compares these data. Population-level or ecological data is sometimes collected across different

time points, making longitudinal ecological studies another possible approach.

- **Follow-up study:** follow-up studies are particularly interested in the trajectories of individuals after a particular event. For example, people who are at risk of violent extremism may be followed up over time to see what kinds of experiences they have, and what the trajectories of these experiences are.
- **Retrospective study:** in some instances it is possible to do a longitudinal study by simply looking back in time at records from a particular group of people in the past. For example, the criminal records of a study cohort in the 10 years before recruitment into the study could be examined along with other administrative data from the same time period.

16.2. Purpose

The primary reason to carry out a longitudinal study rather than a cross-sectional study is to understand not just what the current situation is, but also how things have changed or are going to change over time.

16.3. When to use it

Some of the strengths and limitations of longitudinal designs depend on the particular approach taken. However, there are some points of general relevance to longitudinal approaches:

16.3.1. Why?

The advantages of longitudinal designs include the following:

- It is possible to look at change over time, rather than just what is happening at a single point in time.
- When a longitudinal approach includes person-level follow-up over time (for

example in cohort studies or intervention studies), it may be possible to make causal inferences (i.e. to say that A caused B, rather than just that A is associated with B). The classic criteria for whether a study can provide evidence of a causal relationship are the Bradford-Hill criteria (Pickett and Wilkinson, 2015). The cause coming before the effect in time (temporality) is one of these criteria.

- In longitudinal studies, selecting people at random before the outcome of interest has occurred can help to minimise selection bias.

16.3.2. Why not?

There are also disadvantages to longitudinal designs:

- It is inherently more expensive to collect data at multiple time points rather than just one.
- The study design may be more complex than a single data collection exercise.
- More time is required to get an answer because data collection may take many months. A possible way around this is to set up vignette-based experiments within a survey or other form of data collection.
- People recruited at the start of a study may be lost to follow-up over time (attrition). This means that those left at the end of the study may be different from the original group, introducing potential bias to the effective sample.
- Other things unrelated to the study may also change over time, making interpretation of findings more difficult. For example, if new laws or national policies are implemented after the start of the study period, there may be changes relating to these laws that were unanticipated before the research began.

- Analyses will often be more complicated than in other study approaches. For example, with multiple observations of the same person over time, the data are 'clustered' by person, meaning that the assumption of independent observations that underpins many standard statistical techniques is violated. Analyses suited for repeated measures can be used.
- Things may change over time just because people are being observed – a phenomenon known as the Hawthorne effect (Draper, 2016).
- It is generally very difficult to undertake longitudinal studies with hard-to-reach populations such as truants.
- Longitudinal designs cannot be easily applied to rare events unless using a very large sample, or a sample at higher risk of an event occurring.

16.4. Step-by-step application

The mechanics of conducting a longitudinal study vary hugely (see Stouthamer-Loeber et al., 1992). Focus groups, surveys or interviews could be carried out over time across different longitudinal approaches.

What is important to think about at the planning stage of a longitudinal study (or indeed any study) is the research or evaluation question to be answered (CLOSER, 2017). A clear research question makes it possible to think about the details of the particular type of data to be collected or used. Think about whether a comparison group is needed, and what group that might be.

In broad terms, research is either quantitative (things that can be counted) or qualitative (things that are not counted). For longitudinal quantitative studies (for example, a study looking at a change in the proportion of people becoming radicalised over time between

countries), specific statistical approaches are needed in order to account for the clustering of data across different time points. In general, the following common steps are applied:

1. Define aims, objectives and feasibility
2. Sample the target population
3. Determine the mode and design data-collection instruments
4. Collect data (including ethical review, testing, fieldwork, data preparation and data access)
5. Analyse data.

16.5. Ethics

The research ethics of a longitudinal study can vary depending on the methods used. For research ethical issues relating to specific methods, see the chapters of this document on focus groups, surveys or interviews.

There is one ethical issue that is specific to longitudinal studies. From a methodological point of view, it is important to keep respondents in the study and have enough respondents who have provided information at all points of measurement. However, participation is voluntary and respondents are free to stop providing information at any point. Researchers should not put undue pressure on respondents who have provided information at the first measuring point to also participate at later measuring points.

List of references

CLOSER. 2017. 'The learning hub'. UCL Institute of Education. As of 18 February 2018: <https://learning.closer.ac.uk/study-design/>

Draper, S. 2016. 'The Hawthorne, Pygmalion, Placebo and other effects of expectation: some notes'. As of 18 February 2018: <http://www.psy.gla.ac.uk/~steve/hawth.html>

Nivette, A., Eisner, M. and Ribeaud, D., 2017. 'Developmental Predictors of Violent Extremist Attitudes: A Test of General Strain Theory'. *Journal of research in crime and delinquency*, 54(6), pp.755-790.

Pickett, K.E. and Wilkinson, R.G. 2015. 'Income inequality and health: A causal review'. *Social Science and Medicine*, 128:316–326.

Stouthamer-Loeber, M., Van Kammen, W. and Loeber, R., 1992. 'The nuts and bolts of implementing large-scale longitudinal studies'. *Violence and Victims*, 7(1), p.63.

Check Appendix: Additional resources for further reading.



See also:



CHAPTER SEVENTEEN

Meta-analysis

Elta Smith, Ben Baruch and Anke van Gorp

17.1. Brief description

Meta-analysis is one of a set of techniques for synthesising results from previous studies. Meta-analysis is the quantitative (statistical) aggregation of research results. It involves a variety of statistical methods for reviewing and summarising the quantitative results of prior research. Although combinatorial practices vary greatly, the most common practice in meta-analysis is the conversion of study results into an average effect size.

Meta-analyses should only be undertaken under one of two conditions: (1) following a systematic review of comparable interventions and outcomes; or (2) where a study is planned where the same research design is being implemented in different locations (see e.g. Ariel et al., 2016).

17.2. Purpose

Meta-analysis can provide evidence about the effectiveness of an intervention – that is, the extent to which observed outcomes can be attributed to an intervention. Patton

(2014) identifies meta-analysis as one of three principle approaches to assessing the evidence for the effectiveness of an intervention. In addition to meta-analysis, a single-summativ analysis may be used to provide evidence of process and outcomes related to a single intervention. ‘Principles-based’ analysis is used to synthesise the results of a group of diverse interventions that all adhere to the same principles but have adapted those principles to the particular circumstances of the intervention. Compared with conventional literature reviews, evaluation synthesis or meta-evaluation, the most distinctive aspect of meta-analysis is that results of studies are statistically combined. Current examples of meta-analysis come from related fields, such as criminology or public health (e.g. Dowden and Andrews, 2000; DiMaggio and Galea, 2008; Steel et al. 2009), where the use of quantitative and (quasi-) experimental designs is more common.

17.3. When to use it

Meta-analysis can measure the effect of a set of research results where the questions

being asked by the studies are comparable, and where the studies being analysed have all used the same design and measure the same outcomes in a standardised and replicable manner. The research results being analysed should be empirical rather than theoretical, and the analysis should be used to examine the same constructs and relationships. Finally, it should be possible to combine the findings being aggregated in a comparable format, typically an effect size.

Lipsey and Wilson (2001) identify four types of research findings that are suitable for meta-analysis:

- Central tendency research – prevalence rates;
- Pre-post contrasts – growth rates;
- Group contrasts – experimentally created groups (comparing outcomes between treatment and comparison groups) and naturally occurring groups;
- Association between variables – measurement research (validity generalisation) and individual differences research (correlation between personality constructs).

17.3.1. Why?

Meta-analysis can be used to assess evidence across studies in order to establish evidence-based practices. It is particularly useful for:

- Combining the results from many, smaller studies, to synthesise an overall average effect.
- Finding relationships across studies that cannot be identified using other approaches, including studies for which independent tests of significance have been used but there is little evidence of significance from any one set of results;

- Showing the variability of direction and magnitude of effects across many studies (demonstrated by the 'effect size');
- Identifying knowledge gaps where a meta-analysis is desired but cannot be performed (an 'empty synthesis' underpinned by so-called empty reviews'; see Yaffe et al., 2012). It can also support study design by identifying these gaps and the likely sample sizes needed for a meaningful analysis.

17.3.2. Why not?

Meta-analysis has the following drawbacks:

Comparability of studies used in a meta-analysis may be disputed between researchers.

The approach is subject to selection bias, especially where negative and null findings may not be available (e.g. because significant findings are more likely to be published than non-significant findings). This is why it is important to base meta-analyses on systematic reviews – even then the 'file drawer problem' means that the available studies may not represent the whole picture. The 'reproducibility crisis' in psychology and other subjects had highlighted just how problematic publication biases may be (Ioannidis, 2005; Lösel, 2017; Pridemore et al., 2018)

It is a labour-intensive approach that may not yield results if it results in an 'empty synthesis' due to a lack of studies that can be analysed in this way.

17.4. Step-by-step application

While the use of this method may currently be limited in the area of P/CVE, it will be increasingly useful as the number of robust and rigorous P/CVE evaluations grows. The steps required to use this method are detailed below. They are summarised from Lipsey and Wilson (2001).

1. Identify the relationship to analyse

The first step in any meta-analysis is to identify the relationship to examine. This relationship may be highly abstract, such as the overall effectiveness of different de-radicalisation programmes.

The relationship being examined may also be narrower – it will be easier to demonstrate comparability between studies that represent direct replications e.g. using exactly the same research design and outcomes as the original study, but in a new location, with a different population, and so on. For example, studies that examine effectiveness of EXIT-type interventions that facilitate the withdrawal of individuals from violent and radicalised groups may be more readily comparable because they have simply defined and quantifiable targets (in this case the number of individuals withdrawing from such groups). However, this may not be true if the implementation of EXIT varies between contexts – one has to be sure that the ‘active ingredients’ of an intervention are the same in different locations.

2. Define the boundaries

The studies to be included in the meta-analysis need to be identified, and explicit inclusion and exclusion criteria for sources need to be identified. These criteria may be based, for example, on a particular definition of terrorism or violent radicalisation. Inclusion and exclusion criteria should be developed iteratively through the literature review. Decisions will need to be taken regarding whether to accept low-quality studies for the analysis. Being too restrictive in terms of quality may reduce generalisability, while being too inclusive may weaken confidence in findings.

3. Calculate the effect size

Meta-analysis is dependent on the ability to establish a standardised, numeric, and therefore comparable ‘effect size’ for the selected research findings. There are many different ways in which the effect size can be calculated, including standardised mean differences, odds-ratios and correlation coefficients. Proportions and standardised gain scores can also be used but are not considered to be as strong as the other approaches. One effect size should be calculated for each study or sub-sample within a study that will form part of the meta-analysis.

4. Weight by inverse variance (or by sample size)

Studies included in the meta-analysis should be weighted so that larger (i.e. more precise) studies carry more weight than smaller studies. A simple way to do this is to weight each effect size by its sample size. A better approach is to weight by inverse variance. For this, a standard error must be calculated for the effect size, which generates confidence intervals; smaller standard errors indicate a more precise effect size. Transformations may be undertaken to standardise the effect size to facilitate calculations of the inverse variance weight (e.g. to correct for sample size bias). Adjustments may be needed to correct for measurement unreliability, range restrictions (in order to achieve an unrestricted standard deviation) and normalise the underlying distribution. Outliers should be removed or adjusted to a less extreme value as these can have a disproportionate influence on the analysis.

5. Conduct the meta-analysis

Each study will have an associated effect size and weight. First, multiply the effect size (ES) by the weight (w) for all studies.

Then sum the weights for all studies ($\sum w$) and sum the multiplied effect sizes and weights ($\sum ES \cdot w$) for all studies. Divide the sum of weight * effect size by the sum of all the weights.

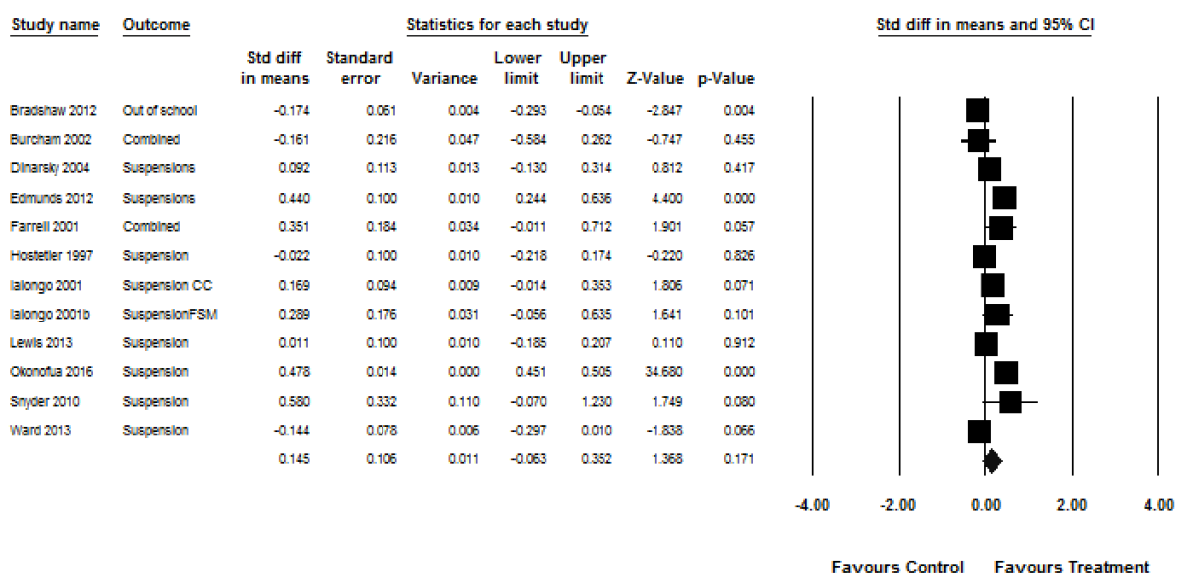
Calculate the standard error of the mean by taking the square root of 1 divided by the sum of the weights. Conduct a Z-test for the mean effect size. Estimate the 95 per cent confidence interval.

Homogeneity analysis should then be used to test whether it is a reasonable assumption that all of the effect sizes are estimating the same population mean. If homogeneity is rejected, the distribution of effect sizes is assumed to be heterogeneous. Heterogeneity can represent a challenge to the basis for combining studies in a meta-analysis. Therefore, if the effect size is heterogeneous, analyse the fixed or random effects of the heterogeneous distributions.

Finally, the effect size results must be interpreted. There are several 'rules-of-thumb' that can be used to interpret the standardised mean difference effect size, the correlation coefficient, and the odds-ratio in terms of whether they are 'small' or 'large'. But, such 'rules-of-thumb' do not account for the context of an intervention, or what is substantively meaningful in terms of difference between groups. For example, a small effect may be highly meaningful for certain types of intervention (e.g. those that require few resources and/or impose little on participants). Small effects may also be important for difficult problems.

As an example of one output from a meta-analysis, we include a 'forest plot' from a recently published systematic review and meta-analysis (Valdebenito et al. 2018). The figure shows the result from each study, along with an overall average effect.

Figure 3: example 'forest plot' from published meta-analysis on school exclusion



Source: Valdebenito et al. (2018:69).

17.5. Ethics

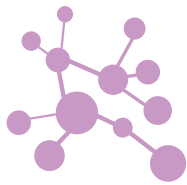
Selection bias might also lead to ethical issues: in selecting what to focus on, a bias might be introduced or a spurious association might be found. For instance, police in the US conducted drug searches based on the assumption that Latinos and Afro-Americans would be overrepresented among drug users (Levine et al. 2010). Because Latinos and Afro-Americans were more often searched for drugs this led to more confidence in the hypothesis; however, sociological research shows that drug use is equally prevalent in other groups that were not searched as often. Searching for evidence of a hypothesised association can immediately create a bias. It is also necessary to actively search for disconfirming evidence.

In most cases a meta-analysis will not use any personal data. If personal data are used, evaluators need to adhere to data protection regulations. For more information, see chapters on cross-sectional data analysis, data mining and descriptive statistics.

List of references

- Ariel, B., Sutherland, A., Henstock, D., Young, J., Drover, P., Sykes, J., Megicks, S. and Henderson, R. 2016. 'Wearing body cameras increases assaults against officers and does not reduce police use of force: Results from a global multi-site experiment', *European Journal of Criminology*, 13(6): 744-755.
- DiMaggio, C. and Galea, S. 2008. 'The Behavioural Consequences of Terrorism: A Meta-Analysis'. *Academic Emergency Medicine*, 13(5):559–566.
- Dowden, C. and Andrews, D. 2000. 'Effective correctional treatment and violent offending: a meta-analysis'. *Canadian Journal of Criminology*, 42(4):449–467.
- Ioannidis, J.P.A. 2005. 'Why most published research results are false'. *Plos Medicine*, 2(8): e124.
- Levine, H.G., Gettman, J.B. and Siegel, L., 2010. *Targeting Blacks for Marijuana: Possession Arrests of African Americans in California, 2004-08*. California: Drug Policy Alliance.
- Lösel, F., 2017. 'Evidence comes by replication, but needs differentiation: the reproducibility issue in science and its relevance for criminology'. *Journal of Experimental Criminology*, pp.1-22.
- Lipsey, M.W. and Wilson, D.B. 2001. *Practical meta-analysis, Applied social research methods series (vol. 49)*. Thousand Oaks, CA: Sage.
- Patton, M.Q. 2014. *Evaluation Flash Cards: Embedding Evaluative Thinking in Organizational Culture*. St. Paul, MN: Otto Bremer Foundation.
- Pridemore, W.A., Makel, M.C. and Plucker, J.A., 2018. 'Replication in criminology and the social sciences'. *Annual Review of Criminology*, 1, pp.19-38.
- Steel, Z., Chey, T., Silove, D., Marnane, C., Bryant, R.A. and van Ommeren, M. 2009. 'Association of Torture and Other Potentially Traumatic Events with Mental Health Outcomes Among Populations Exposed to Mass Conflict and Displacement: A Systematic Review and Meta-analysis'. *JAMA*, 302(5):537–549.
- Yaffe, J., Montgomery, P., Hopewell, S. and Shepard, L.D., 2012. 'Empty reviews: a description and consideration of Cochrane systematic reviews with no included studies'. *PLoS One*, 7(5), p.e36626.

Check Appendix: Additional resources for further reading.



CHAPTER EIGHTEEN

Network analysis

Gavin Cochrane and Anke van Gorp

See also:



Example(s) of use

Network analysis was used in a study which explored the structure and content of criminal connections and tested a belief that contemporary organized crime groups take advantage of globalization by opening their bases in distant territories (Varese, 2012). The study examined police data pertaining to an attempt by a Russian mafia group to open a branch in Italy and found that the group was in fact forced to move abroad.

18.1. Brief description

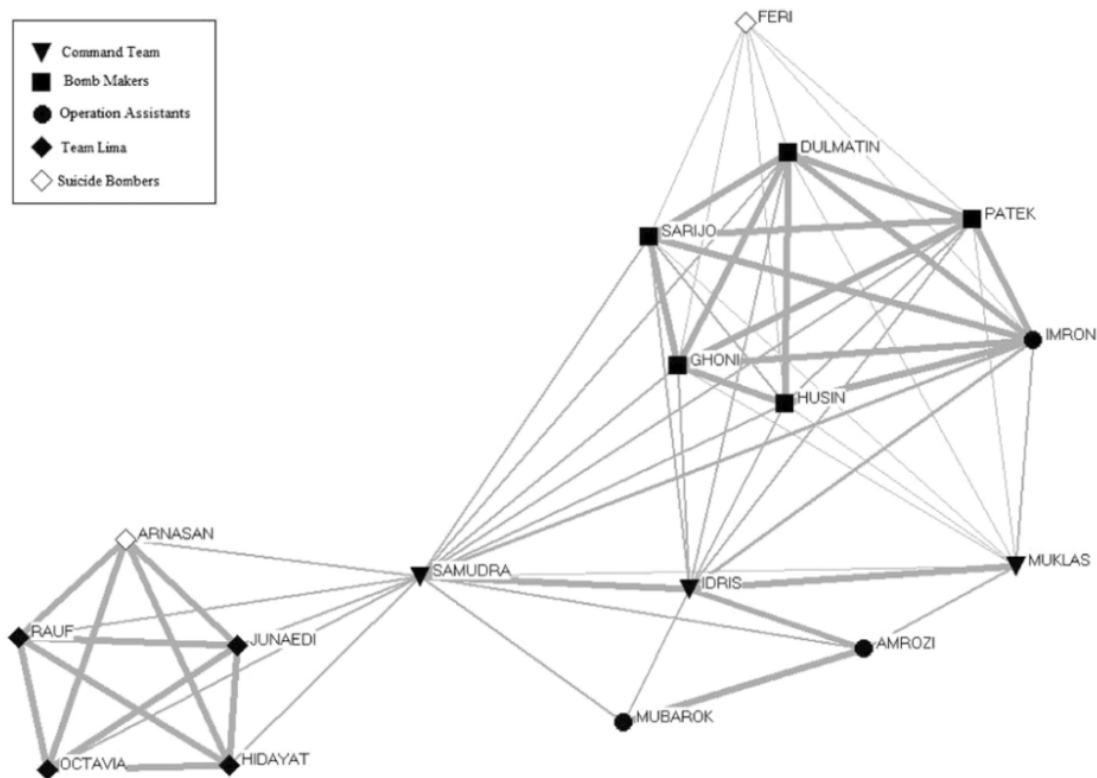
Network analysis is the process of mapping and measuring relationships and flows between individuals/groups. It has been used to examine the social structures and dynamics of both radical groups and organisations that work to prevent and counter violent extremism. Network analysis is used to describe, both numerically and graphically, the structure of and interactions within a network and can be used to identify key actors.

Networks are commonly presented as a pattern of connections between nodes, which can represent a variety of actors in a given network (e.g. individuals in a terrorist cell, European

institutions working in P/CVE, collaboration between countries in funding P/CVE activities). Analysing the structure of such networks can be used to develop perspectives, models and paradigms on interactions between actors in a network and how influential different actors can be within a network.

In particular, network analysis can be undertaken to provide answers to the following questions:

- Which individuals/groups are central in the network?
- Which individuals/groups are peripheral to the network?
- Which connections are the most crucial in the functioning of the network?

Figure 4. Jemaah Islamiyah Graph – Bali Operation October 6–11, 2002

Source: Koschade (2006).

18.2. Purpose

A network approach provides a systemic perspective on social or organisational structures, taking into account the dynamics of the system as a whole as well as the relative position of individual actors/groups within a network, thus affording greater analytical detail. The value of network analysis as an approach to P/CVE compared to other methodologies is 'its focus on the value of the network structure rather than the characteristics of the individual' (Ressler 2006:2). Network analysis is also commonly used to construct diagrams of known connections within radical groups (see Figure 4).

18.3. When to use it

Researchers in a number of fields have used network analysis to understand and explain social structures. Network analysis can be particularly useful in studying radical groups and organisations that work to prevent and counter violent extremism. It offers the opportunity to measure the effect of integration, cohesion, and partnership through, for example, assessing whether interventions have resulted in changes to links between individuals or between individuals and institutions (see e.g. Paluck, 2011).

18.3.1. Why?

According to Hunt (2010), network analysis can be a particularly useful tool in:

- Illustrating a complex system: network analysis can provide a useful way of depicting the interactions between actors, visually representing communication channels and how information, ideas or knowledge are shared between individuals/groups (e.g. illustrating how participants in a P/CVE intervention, as well as their partners, families and communities, interact with one another).
- Creating an understanding of the relationships between individuals/groups: network analysis can provide an empirical framework to measure social interaction between individuals/groups over time (e.g. measuring whether and how the participant's interactions with his or her network changes over the course of a P/CVE intervention).
- Identifying problems with the flow or existence of a network: network analysis can also help to identify where gaps or bottlenecks in interactions may appear, or vulnerabilities in a network (e.g. in the case of a P/CVE intervention that works by signposting participants to public services such as social and educational services, network analysis can be used to identify where services may be needed).

18.3.2. Why not?

While network analysis can provide a visual representation of social structures, it does not offer normative assessments on the way networks should be or why relationships/interactions exist between actors. This is also a function of how network connections are measured (e.g. telephone calls, texts, emails, face-to-face interactions).

Network analysis also requires an understanding of the contextual background of the social structures being examined. This element is crucial in designing appropriate mechanisms for data collection and defining the boundaries of the network. For example, disconnected units and asymmetries in information on radical groups make it difficult to ensure that the network is capturing all relevant actors.

Network analysis is not an appropriate method for suggesting how things ought to work in a network and does not capture how interactions between nodes occur on a day-to-day basis. Even a longitudinal network study aimed at the evolution of a network over time only offers snapshots, and may miss many configurations of the network in-between.

18.4. Step-by-step application

The steps required to use this method to assess networks in the area of P/CVE are detailed below.

1. Define the boundaries

A preliminary step in conducting a network analysis is to determine the population under investigation by identifying the relevant actors for inclusion and defining the boundaries of the network. This can be done using either a realist or nominalist framework. The former focuses on predetermined groupings set by actors (e.g. members of a radical group), while the latter is constructed by the researcher (e.g. everyone interacting with members of a radical group in a given timeframe or location).

2. Collect data

The collection of data for network analysis involves developing a picture of the connections between individuals/groups. A number of methods can be used to

Table 3. Reported working relationships

	Choice:			
Chooser:	Person A	Person B	Person C	Person D
Person A	---	0	1	1
Person B	1	---	1	0
Person C	1	1	---	1
Person D	0	0	1	---

Source: Hunt (2010).

compile data on the connections between actors in the network being examined, including surveys, questionnaires, interviews and observations. The two types of data most commonly collected are structural data on the strength/frequency of connections between actors (e.g. how many times A calls B), and composition data on the various attributes of actors in the network. This can include attribute data on the units of study in the network, *nodes* (e.g. age, gender, geographic location), or on connections between nodes, known as *edges* (e.g. those giving orders versus those receiving orders).

3. Design the network

Once structural and composition data has been collected, the information can be displayed either in matrix (table) format or graphically. There are a number of software packages that can be used for this, including UCINET, Pajek and Gephi.

Matrices are used to display the similarities and differences between actors in a network in a rectangular array, whereby the rows and columns of the table represent individual actors and the cells of the matrix represent the strength of their relationship (see Table 3).

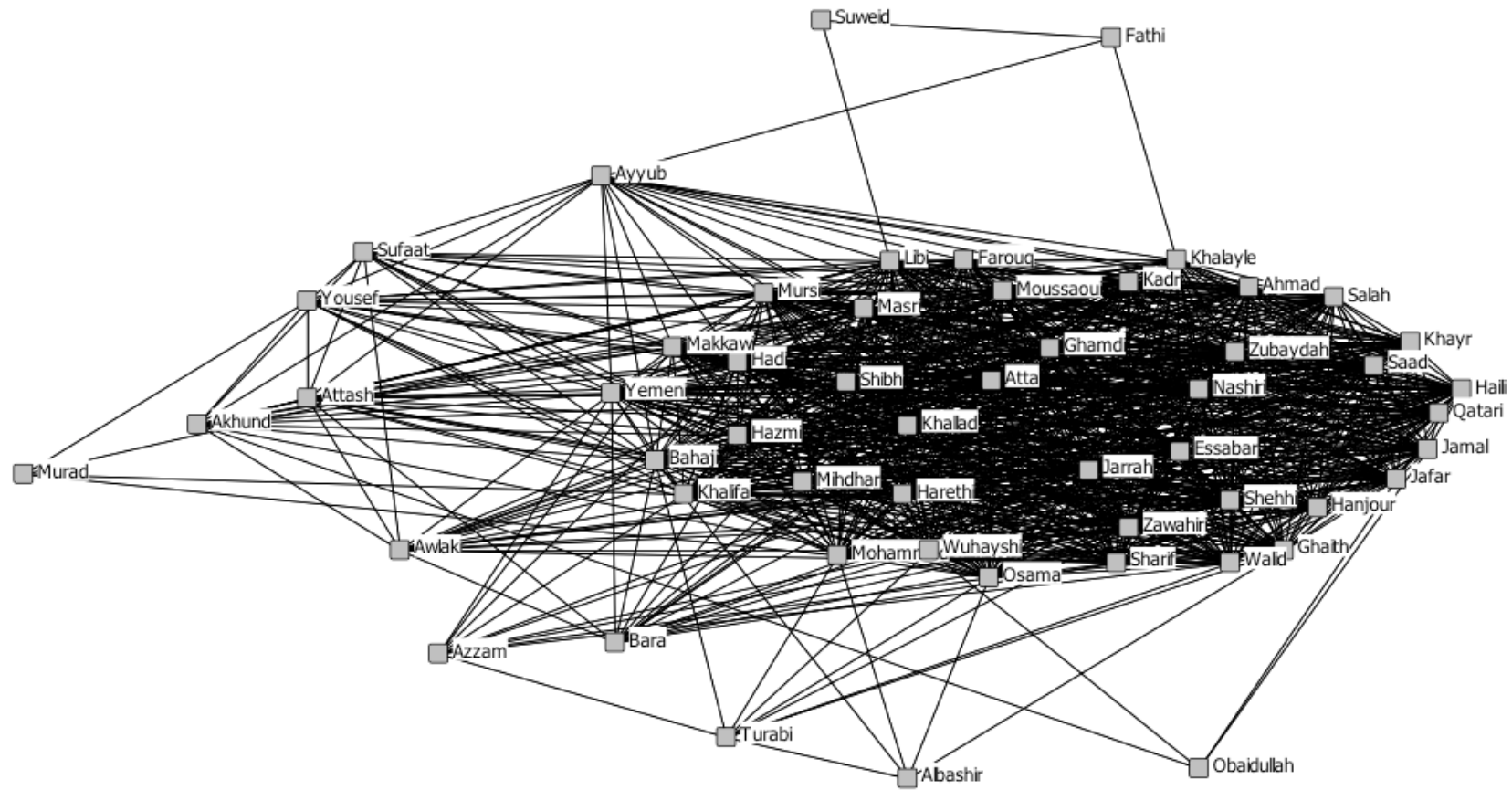
This information can also be displayed graphically, to present a visual representation of the network. These diagrams are composed of points which represent the actors in a network, called *nodes*, and lines connecting them to represent their connections, called *edges* (see Figure 5).

4. Analyse the network

Once the data is represented, in either matrix or graphical form, it can be analysed to determine the structure of the network and establish which actors have the most influence in the network. Centrality measures are often used to analyse a network. The most common include:

- Degree centrality, which measures the number of relationships between actors to establish who has the most connections (the actor who is 'most connected' is typically someone important to the network, regardless of seniority).
- Betweenness centrality, which examines the level of control in relationships through examining nodes that join clusters (subnetworks) rather than nodes that lie inside a community. It is calculated by measuring the number

Figure 5. Al-Qaeda Terrorist Network, members and links between members



Source: Wu et al. (2014).

of shortest paths from all nodes to all others that pass through that node.

- Closeness centrality, which examines the familiarity within relationships to determine how fast information can spread from a given node to all reachable nodes. It is measured by calculating the average path needed to reach all nodes, excluding those to which no path exists.
- Eigenvector centrality, which takes into account differences in the strength of relationships across a network. Degree centrality calculates the number of relationships a node has but does not account for the fact that these connections may be relatively isolated. Eigenvector centrality gives relative scores to connections, with a greater weighting assigned to connections with more influential nodes.

18.5. Ethics

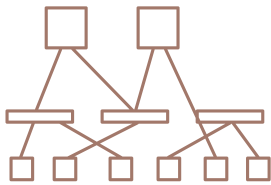
Informed consent is a crucial ethical concern. In a network analysis, information about non-participants can be gathered if participants who are linked to them are involved in the research through asking about them. Even if individuals in a network do not want to participate in the research and provide information, they will still be in the network analysis. This makes informed consent problematic as there is no real possibility of refusing to provide data, because others can provide data about an individual and their relationships. An individual's refusal to participate would only lead to no information being collected about them if a large part of the research population also refused to participate, making a network analysis impossible. Borgatti and Molina (2003) claim that this is the nature of network analysis and is acceptable, but the interests of individuals who are not participants should be taken into account.

For discussion of ethical issues surrounding methods that can be used during network analysis (such as surveys, questionnaires, interviews and observations), see the chapters covering those methods.

List of references

- Borgatti, S.P. and Molina, J.L. 2003. 'Ethical and Strategic Issues in Organizational Social Network Analysis'. *The Journal of Applied Behavioral Science*, 39(3):337–349.
- Hunt, T. 2010. *The Power of Social Networking: Using the Whuffie Factor to Build Your Business*. Crown Publishing.
- Koschade, S. 2006. 'A social network analysis of Jemaah Islamiyah: The applications to counterterrorism and intelligence'. *Studies in Conflict and Terrorism*, 29(6):559–575.
- Paluck, E.L., 2011. 'Peer pressure against prejudice: A high school field experiment examining social network change'. *Journal of Experimental Social Psychology*, 47(2), pp.350-358.
- Ressler, S. 2006. 'Social network analysis as an approach to combat terrorism: past, present, and future research'. *Homeland Security Affairs*, 2(2):1–10.
- Varese, F., 2012. 'The structure and the content of criminal connections: The Russian Mafia in Italy'. *European sociological review*, 29(5), pp.899-909.
- Wu, E. Carleton, R. and Davies, G. 2014. 'Discovering bin-Laden's Replacement in al-Qaeda, Using Social Network Analysis: A Methodological Investigation'. *Perspectives on Terrorism*, 8(1):57–73.

Check Appendix: Additional resources for further reading.



CHAPTER NINETEEN

Objectives and options analysis

Kate Cox and Anke van Gorp

See also:



19.1. Brief description

Objectives analysis (also known as situational analysis or 'problem and objective tree analysis') is a methodological approach whereby problems are identified and then converted into project goals. The two key outputs of objectives analysis are the 'problem tree' and 'solution tree'. A problem tree provides an overview of all of the known causes and effects of an identified problem (e.g. social segregation, social welfare issues, low resilience, etc.), while a solution tree replaces the problem statements with positive objectives and provides an overview of interventions (e.g. help with access to housing, training, employment, increasing empowerment, etc.) to address the core issue.

Options analysis is an approach that considers how best to reach project objectives. It involves a systematic assessment of all possible options available for achieving the aims of a project and is implemented once

project objectives have already been defined. During the analysis, each option is described, its likely impacts are assessed, and it is screened against criteria such as effectiveness, efficiency and consistency in order to create a shortlist of the most promising options. This options shortlist is then subjected to a more detailed impact analysis and comparison.

19.2. Purpose

Objectives analysis and options analysis can both be used as a basis for planning a P/CVE project, but they can also help in drawing recommendations from an evaluation for the next generation of projects. Objectives analysis assesses the causes of a specific problem (e.g. high rate of radicalised adolescents) and determines how they can be addressed (e.g. strengthening identity, empowerment, rehabilitation, etc.). It provides an outline of the project plan, including the activities that need to be undertaken, the goal and the desired outcomes of the project.

Options analysis ensures that selected options support project objectives and contribute to successful delivery. By analysing project goals and identifying alternative means of achieving them, options analysis can lead to a better understanding of the project that facilitates effective implementation and evaluation.

19.3. When to use it

Options analysis and objectives analysis are both carried out during the project planning process. They can be used during the data gathering phase to identify issues that can later be explored through other methods (e.g. surveys or focus groups), or to explore issues already identified through literature review. Options analysis is often implemented after the objectives analysis process has already defined and assessed project goals. In evaluation, options analysis can be used once data is collected in a process of drawing conclusions and recommendations for the future.

19.3.1. Why?

There are a number of ways in which options analysis and objectives analysis can be valuable, including:

- To break down the problem into manageable and definable components (options and objectives analysis);
- To help establish whether further evidence and/or resources are required to create a suitable solution (objectives and options analysis);
- To facilitate a better understanding of problems and their causes (objectives analysis);
- To identify the relevant actors and processes at each stage (objectives analysis);

- To allow researchers to identify and explore the most cost-effective means of realising project objectives (options analysis).

19.3.2. Why not?

These types of analysis also have a number of drawbacks, including the following:

- Listing all possible solutions at an early planning stage can hinder objective and open-minded analysis (objectives analysis).
- The problem tree gives no indication of the magnitude of the problem – all problems are seen as equally important (objectives analysis).
- It is not possible for every problem to be rephrased as a solution; in other words, not every problem can be subjected to objectives analysis (objectives analysis).
- The analysis does not take resource and time constraints into account when identifying solutions to problems (objectives analysis).
- A linear (means and end) representation of a problem and solution may be inappropriate for complex cases unless assumptions are identified and tested (objectives analysis).
- Too many or too few options might make a decision more difficult to take (option analysis).
- If done badly, the analysis undermines the credibility of the study and its usefulness for political decision making (option analysis).
- Creating a problem tree and an options analysis of a programme can be problematic since it may ignore some of the underlying complexity and social reality of the intervention.
- A linear (means and end) representation of a problem and solution may be inappropriate for complex cases unless

assumptions are identified and tested. It is therefore necessary to transparently present the methodology and assumptions.

- It may be that not all relevant options are identified.

19.4. Step-by-step application

This section provides a step-by-step guide to objectives and options analysis. It is best carried out in a small group of about six to eight people involved in project planning or evaluation.

Objectives analysis steps

Objectives analysis involves a three-phase process that first analyses problems and then converts these into desired solutions, before selecting a preferred intervention strategy.

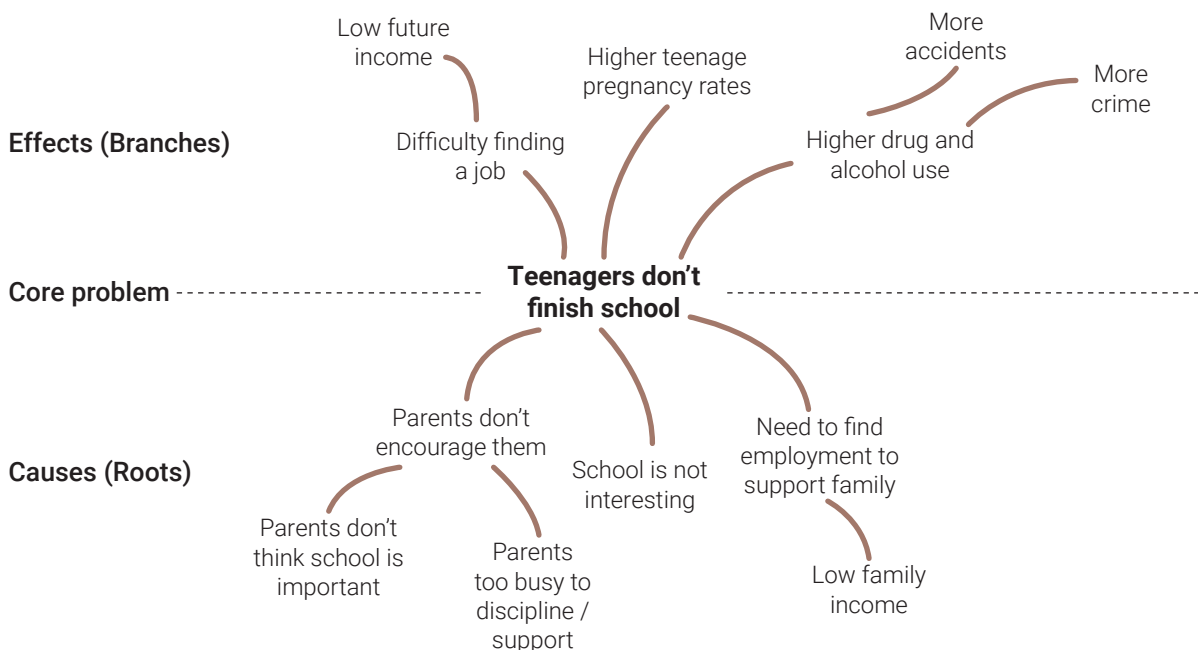
1. Develop a problem tree
 - Identify the problem that the project seeks to address.
 - Write down the core problem (the 'trunk') on a post-it and then stick it to the middle of a wall.

- Identify and apply lessons from previous projects to help define the core problem.
- If there appears to be more than one core problem, develop a problem tree for each one.

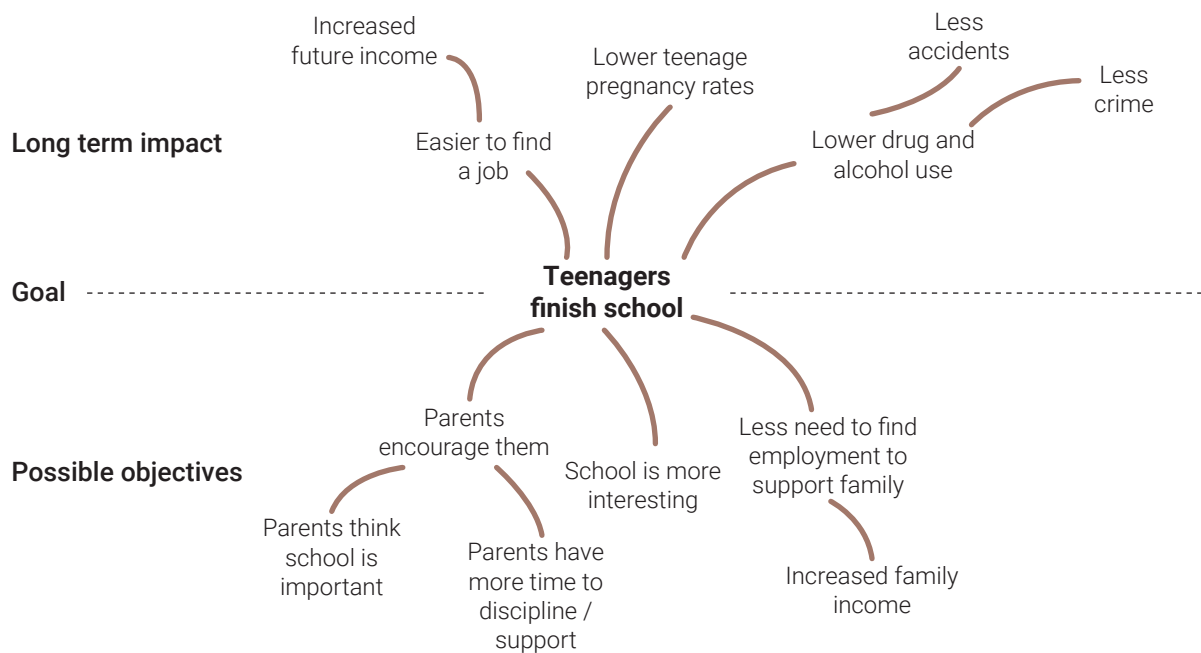
- Determine the causes and effects of the problem.

- Participants should collectively brainstorm the causes ('roots') of the problem with a facilitator writing each suggestion on a post-it. The post-its should then be stuck on a wall for participants to analyse and re-order, before linking to effects ('branches').
- Alternatively, participants could work through the cause and effect on a sequential basis, starting with the core problem. Participants should continue this process until they cannot identify any further underlying causes.

Figure 6. Problem tree



Source: tools4dev (2014).

Figure 7. Objectives tree

Source: tools4dev (2014).

2. Develop an objectives tree

The objectives tree is another key tool at the project planner's disposal and is closely linked to the problem tree. There are five basic steps in developing an objectives tree:

- Rephrase each problem into a positive solution in an objectives tree diagram. For example, 'lack of knowledge' would become 'increased knowledge'. When the problem cannot easily be converted into an objective, the problem should be reconsidered and more clearly articulated.
- Consult with stakeholders when reviewing objectives to ensure that they are desirable and achievable in an acceptable timeframe.
- Modify or remove objectives that do not meet these conditions.
- Examine the 'means–end' relationships derived to check the validity, logic and

completeness of the diagram, and modify as necessary.

- Divide the objectives tree into clusters, and select the clusters that will be included in the intervention.

3. Select the preferred intervention

- Focus the intervention on a preferred strategy, taking into consideration available project funding, as well as time and relevance.
- Use the information captured in the problem tree and solution tree to develop a logic model, which provides a summary of key elements of the project and monitoring plan.

Options analysis steps

After analysing and setting out objectives for the intervention, the next step is to consider how these objectives can be achieved. There are probably different ways to achieve them,

each with certain advantages, disadvantages and its own unique requirements. It is worth considering which is the best path to take, and options analysis facilitates that. Options analysis should be carried out after objectives analysis. There are three basic steps:

1. Identify options for each project
 - Describe a baseline scenario ('do nothing' option): identify the 'business as usual' projection that involves incurring operational and maintenance costs within already existing infrastructures.
 - Define a 'do minimum' option for the project: describe a scenario that requires minimum effort and expense beyond existing operational and maintenance costs.
 - Explore 'do something' options for the project: identify options that involve a level of investment that depends on project objectives.
2. Conduct a feasibility analysis
 - For each option, determine the following:
 - Inclusions, exclusions, assumptions, constraints
 - User requirements
 - Issues to be resolved
 - Deliverables.
 - Summarise the analysis and highlight the most promising solution for the project in a report which should also include the results from previous steps.
3. Finalise option selection (for major projects and large-scale interventions only)
 - Perform cost-effectiveness analysis: compare alternatives with the aim of selecting an option that minimises costs and/or maximises the output level.
 - Evaluate the economic impact: use impact analysis as a way to select the best alternative for the project design.

Identify and evaluate the foreseeable impact of every option on the economic status of the project. To do this, focus on using high-level economic indicators and forecast their influence on the project environment.

- Finalise choice of option: summarise all the steps taken, confirm whether the analysis has demonstrated that feasible alternative options have been adequately considered, and verify that the best option has been selected for the project design.

19.5. Ethics

When conducting objectives and options analyses, the following ethical points should be considered:

- Ethical issues may arise depending on the methods used to gather data. See chapters on interviews, surveys, focus groups, literature review, case studies, etc.
- If this type of analysis is used in an ex-post evaluation, stakeholders might be affected by what they have experienced and learned during the intervention and therefore might be unable to correctly reconstruct what assumptions and ideas they had during the design of the intervention. This is known as hindsight bias (see for example Roese and Vohs, 2012).

List of references

- Haringey Council. n.d. *Haringey Prevent situational research: Executive Summary*. As of 2 September 2015: http://applications.haringey.gov.uk/haringey_prevent_situational_research_exec.pdf
- Roese, N. J. and Vohs, K.D. 2012. 'Hindsight Bias'. *Perspectives on Psychological Science*, 7(5):411–426.

tools4dev. 2014. 'Resources: How To Guides. How to design a new program'. As of 18 February 2018:
<http://www.tools4dev.org/resources/how-to-design-a-new-program>

Check Appendix: Additional resources for further reading.



CHAPTER TWENTY

Observation techniques/ ethnography

Marie-Louise Henham, Megan Sim and Anke van Gorp

See also:



Example(s) of use

Observational techniques have been used to understand how policies operate in the education, health, justice and criminal systems.

Cultural Interactive is a NGO in Germany that works with young people to prevent radicalisation and carry out de-radicalisation among those who are vulnerable to violent right-wing extremism, ethno-nationalism or religious fundamentalism. As part of its work, Cultural Interactive offers advanced training in methodology for youth workers, including the Locally Embedded Derad-Training (LocalDerad), which delivers practical tools for local assessment, immersed observation, narrative interviewing and self-evaluation (Radicalisation Awareness Network, 2014).

20.1. Brief description

Observation techniques are a key element of action research which allow a phenomenon to be studied in its natural setting. The two main approaches are non-participant observation and participant observation.

Non-participant observation is a method of data collection that is commonly used in case study research and has historically been

associated with the social and behavioural sciences. Non-participant observers can have varying levels of separation from their subject. Furthermore, a non-participant observer can either be completely visible and acknowledged by the subjects or be concealed from view. For example, some observers have no contact with their subject and will use cameras or one-way mirrors to observe their subject. For instance, in a P/CVE intervention that employs

one-to-one contact between participants and mentors, the evaluators could observe the mentoring session by being in the room while it is ongoing, through a one-way mirror, or by watching a video recording of the session afterwards.

Participant observation has firm roots in anthropological studies such as ethnography. Participant observation takes place when an evaluator engages directly with the subject of their observations (e.g. joining a population or organisation) and participates in the activities that are being studied. For instance, in a P/CVE intervention that conducts group sessions with potentially radicalised individuals, the evaluator-observer could undertake sufficient training to participate as a facilitator of the group session. It is key that throughout the process the participant observer fulfils two objectives: to participate and experience, and at the same time to interpret and analyse.

20.2. Purpose

Observation can give an evaluator first-hand experience of activities and how actors engage with them. It can provide details about which actors interact with whom and how they do so, how long activities take, and what non-verbal expressions of feeling people convey. Observation can also be used to examine descriptions already provided by informants and validate findings. Furthermore, observations may provide an opportunity for evaluators to access information which actors may be unwilling or unable to share in an interview. For instance, evaluators may observe a participant's facial expressions and body language during an intervention, or the nature of their interaction with intervention specialists, to assess their level of satisfaction with the programme and its facilitators.

An evaluator using non-participant observation will observe phenomena at a distance and will

aim to garner an understanding of events and behaviours in their natural context without participating directly in the activities being observed.

Through participant observation the evaluator is able to gain a fuller appreciation of an activity and its importance to the actors involved. Furthermore, through being involved in the action the evaluator is able to experience the activity first-hand and may be able to gain additional insights.

20.3. When to use it

20.3.1. Why?

Some strengths of observational methods include the following:

- Observational methods can produce very rich and detailed information and can be highly effective during the implementation of an intervention. Using observation in this context, an evaluator can gain insight into the observable output of the intervention and the influence of context and is sensitive of the viewpoints of the key actors and the beneficiaries (European Commission, 2013).
- If a P/CVE intervention is not proving to be effective, observational methods may help provide a better understanding what aspects of the intervention might not be working well. Observational techniques can be particularly useful in understanding how an intervention is understood at a decision-making level in comparison to how those in the field understand it.
- If there is difficulty in gaining access to a certain field (e.g. due to organisational conflict) or if a population is resistant to more formulaic methods of data collection, this may be the only accepted and appropriate technique.

- This methodology can capture unexpected data as the evaluator is not entering the field with a pre-determined framework with which to collect data.
- Participant observation enables an evaluator to see what day-to-day life is like when people are attending an intervention programme, and allows for much greater depth about a person's life, background and what interventions really mean to those participating in them.
- Observational methods can also provide valuable evidence of how an organisational process can affect behaviour.

20.3.2. Why not?

There are also various disadvantages to the use of observational techniques. For example:

- These techniques can be time-consuming and can generate a large volume of data that is rich in detail and demands careful and lengthy processing and analysis. Thus the techniques can be very expensive.
- Observations are usually limited to a small number of settings.
- These methods place a high demand on the evaluator to be able to effectively absorb and then consider the behaviours of the observed subject.
- It may take time for participants to accept the evaluator and behave in a normal way. For instance, in a group P/CVE intervention where a strong group dynamic has been developed, introducing a new observer may affect participants' levels of trust and change their behaviour.
- In participant observation the characteristics of an evaluator (gender, ethnicity, religion, social status, etc.) may affect the results because data collected is subject to the evaluator's perception (subjectivity of results), and because

data collected will be limited to observed individuals (or groups) and situations. It is also worth noting that the evaluator may not fully appreciate how their individual views may affect observation and analysis. Reliable and robust data depends on a professional and experienced observer-analyst.

20.4. Step-by-step application

1. Choosing the site for observation

The sites where observation is going to take place need to be discussed with key stakeholders and access to these sites needs to be agreed. For example, to evaluate a prison-based programme the observer would have to complete the necessary paperwork to access the prison. Furthermore the evaluator needs to familiarise themselves with the setting and context (this could involve mapping out the setting and networks that operate within it). The evaluator also needs to select key stakeholders (e.g. the frontline staff and participants of a P/CVE intervention) and must allow time to establish a trusting rapport with them if conducting participant observation.

2. Observation

As outlined by Kawulich (2005), observation can be conducted in three ways:

1. Descriptive observation involves observation of every detail possible. This can include notes on the setting, actors, activities and interactions.
2. Focused observation often follows a set of interviews where an evaluator may already have insight and have made decisions about what aspects to focus on during an observation.

3. Selective observation is considered to be the most systematic approach and involves an evaluator focusing on various activities and understanding what distinguishes one from another.

In non-participant observation, the observer carefully observes the phenomena and takes particular care not to disturb the scene and the key actors involved. The observer will need to take detailed notes and audio recordings. Notes can be taken during the observation and afterwards. It is advised that a wider analysis of the context is undertaken to better understand factors which may influence actors in a particular setting.

For participant observation, the observer needs to decide on the degree to which they are going to participate in the subject of their observation. Whilst there is no set timeframe for participant observation to be undertaken, it is recommended that an observer is engaged with their subject and context long enough to understand the environmental and cultural context and to build trust with 'regular' actors. With regard to taking notes of their observations, it is recommended that a participant observer takes breaks from being immersed in the field and continuously updates a field diary. Making notes in this way will make it necessary to remember things accurately and decide on what events and details are important to report. Field notes may include quotations, activities and their order, descriptions, background information and the observer's own thoughts, with information such as time, date and setting clearly noted.

3. Data analysis

Notes must be anonymised and a model can be developed whereby the data can be organised according to a protocol.

This may involve drawing data points together using key words or by actor type or activity. Through using a framework, narratives should form and the evaluator should aim to construct an impression of the scene which has been observed. One way of developing the analysis is to first gather a report of what happened in terms of activities, then what was observed, followed by what was recorded in the field and lastly what can be derived from the observer's personal notes.

20.5. Ethics

An obvious issue with observational research is the privacy of the people concerned (for instance, both the clients and the frontline workers being observed). This issue is more pronounced if people are not aware they are being observed – although it may be rare that such an approach would be possible in this context. Data should be gathered with informed consent, which means that the following ethical procedures should be adhered to:

- Clarify the purpose of the study and the anticipated audiences for the data gathered at the outset.
- Seek informed consent from each person observed.
- Give the people who have been observed the opportunity to see how the observations are reported in the context of research and allow them to respond to the findings.
- Make sure not to note and report any data that is unnecessary for the study and might be considered confidential by the people who are being observed.
- Note that directly attributed comments or observations in reports require the explicit permission of the person being observed.

- Use pseudonyms in referring to individuals and institutions. While this does not guarantee anonymity, it reduces the likelihood that individuals and institutions will be identifiable.

Disguised or concealed observation is strongly discouraged from an ethical point of view. It should only be used if there is no other way to obtain the information and only with sufficient guarantees with regard to the interests and privacy of the people being observed.

List of references

European Commission, Directorate-General for Regional and Urban Policies. 2013. *Evalsed Sourcebook: Methods and Techniques*. European Commission.

Kawulich, B. 2005. 'Participant Observation as a Data Collection Method'. *Forum: Qualitative Social Research*, 6(2): Art. 43. As of 18 February 2018: <http://www.qualitative-research.net/index.php/fqs/article/view/466/996#g1>

Radicalisation Awareness Network. 2014. *Preventing Radicalisation to Terrorism and Violent Extremism: Strengthening the EU's Response*. *Radicalisation Awareness Network (RAN) Collection. Approaches, lessons learned and practices*. European Commission. As of 18 February 2018: <http://www.euopris.org/?wpdmdl=2917>

Check Appendix: Additional resources for further reading.



See also:



CHAPTER TWENTY-ONE

Policy scientific approach

Megan Sim and Anke van Gorp

21.1. Brief description

A policy scientific approach is one means of conducting a theory-based evaluation. Other approaches and methods include realist evaluation, ToC and contribution analysis, all of which can be employed in similar ways to the policy scientific approach. Theory-based evaluations start from the position that programmes, policies or interventions are based on an underlying programme theory, which articulates the logic of the intervention. Theory-based evaluations therefore assess how inputs (e.g. funding from a local authority) lead to outputs (e.g. training courses for community leaders or regular local community events) that help to achieve the goals of the intervention (e.g. fewer radicalised youths within the community). The policy scientific approach first reconstructs the programme theory and tests whether the theory holds up against the intended or observed results. The approach can then be used to improve the programme theory.

21.2. Purpose

Theories underlying a programme or policy are not always explicit, clearly specified or directly

visible to evaluators. For instance, it may not be clear how an intervention that provides investment for school-based programmes will lead to greater cohesiveness and trust within communities. In such instances, evaluators can use a policy scientific approach to find or reconstruct the programme theory, articulate the programme theory in a testable way, and finally assess the theory.

21.3. When to use it

In evaluation a policy scientific approach can be used prospectively (before the intervention starts), concurrent to the implementation of a P/CVE programme, or after the intervention has been completed.

21.3.1. Why?

The advantages of employing a policy scientific approach include the following:

- The policy scientific approach encourages and stimulates the use of multiple methods (such as literature review and interview data) to validate the reconstruction of the programme theory. In particular, the use of desk reviews may counter any subjectivity

introduced by interview and focus group participants.

- The policy scientific approach makes use of diagrammatic representation of the underlying programme theory, which can facilitate dialogue with stakeholders and others.

21.3.2. Why not?

The policy scientific approach also has a number of drawbacks. For example:

- Using policy scientific approach to reconstruct a programme theory involves a high degree of stakeholder participation, which might be costly and difficult (Vaessen, 2006).
- The applicability of the policy scientific approach is limited to programmes with a common or unified programme theory, and it may not be suitable for programmes with multiple levels and sites, or programmes which face internal conflict (Hansen and Vedung, 2010).

21.4. Step-by-step application

The steps involved in the implementation of a policy scientific approach are outlined below.

1. Identify the behavioural mechanisms that are expected to solve the problem
Reviewing formal and informal documents (e.g. published documents, web pages, emails) and interview data can elicit statements that indicate why it is believed that the policy problem at hand must be solved (e.g. statements such as, 'A terrorist attack is highly likely in the country') and what the goals of the programme or intervention are (e.g. 'to prevent terrorist attacks from occurring'). These statements can also reveal the mechanisms of the programme, or what is believed to make the programme effective. Statements such

as the following are particularly helpful for identifying mechanisms:

- 'It is evident that [this programme] will work because...'
- 'The best way to solve this problem is to implement [this policy]...'
- 'Our experience in the field tells us that...'
- 'We believe that...'

2. Link the behavioural mechanisms with the goals of the policy, programme or intervention under review
3. Reformulate the links as conditional 'if-then' or similar propositions

Examples of such propositions include the following:

- 'If [this intervention] is implemented, then [a certain goal] can be met...'
 - 'More [of this programme] will lead to more of [a particular result being achieved]...'
4. Search for 'warrants' to identify the missing links in or between different propositions

Warrants refer to the 'because' part of an argument, which states that one statement follows from another because of a generally accepted principle. For instance, the statement, 'The nation's counterterrorism strategy will be effective this year,' follows from the statement, 'The strategy has been effective for the past 5 years.' In this case, the warrant may be that 'Past performance is the best predictor of future performance.'

Warrants are often implicit and must be inferred by the evaluator. Argumentational analysis (a tool commonly used in logic and philosophy) can be used to analyse chains of arguments and identify and fill in missing links.

5. Reformulate the warrants as a set of 'if-then' propositions and draw a chart of the links

IT tools that can be used for argument visualisation include Rationale, Belvedere, Cohere and Questmap (Leeuw, 2012).

6. Assess the validity of the set of propositions

This involves assessing the logical consistency and the empirical content of the propositions. The latter consists of analysing the extent to which the theory and the expected or observed impact of the behavioural mechanisms correspond to the scientific evidence on the mechanisms.

7. Evaluate the reconstructed programme theory

This can be done by:

- Comparing and juxtaposing it to different programme theories;
- Testing the programme theory empirically, making use of primary and secondary qualitative and quantitative data;
- Refining the programme theory iteratively through soliciting stakeholder feedback and using multiple data collection techniques and sources;
- Comparing the programme theory to existing reviews and synthesis studies.

- A potential ethical issue specific to theory-based evaluation is the difficulty of reconstructing the programme theory without falling prey to hindsight bias (Roese and Vohs, 2012). Stakeholders might be affected by what they have experienced and learned during the intervention and therefore might not be able to correctly reconstruct the assumptions and ideas they had during the design the intervention under evaluation. Moreover, if they have invested in a project already, the 'sunk cost fallacy' might also cloud stakeholders' judgment (Dobelli, 2013).

List of references

Dobelli, R. 2013. *The art of thinking clearly*. London: Sceptre.

Hansen, M.B. and Vedung, E. 2010. 'Theory-based stakeholder evaluation', *American Journal of Evaluation*, 31(3)295–313.

Leeuw, F. 2012. 'Linking theory-based evaluation and contribution analysis: Three problems and a few solutions'. *Evaluation*, 18(3)348–363.

Roese, N. J. and Vohs, K.D. 2012. 'Hindsight Bias'. *Perspectives on Psychological Science*, 7(5):411–426.

Vaessen, J. 2006. 'Programme theory evaluation, multicriteria decision aid and stakeholder values – a methodological framework', *Evaluation*, 12(4):397–417.

21.5. Ethics

When conducting a policy scientific analysis, the following ethical points should be considered:

- Ethical issues may arise depending on the methods used to gather data. For examples, see chapters on interviews, surveys, focus groups, literature review and case studies.

Check Appendix: Additional resources for further reading.



CHAPTER TWENTY-TWO

Qualitative data analysis and grounded theory

Kate Cox and Anke van Gorp

See also:



Example(s) of use

A doctoral study using a grounded theory approach was recently conducted at the University of Bradford, UK, investigating the UK government's Prevent programme and British Muslims living in the North of England. The study analysed the process of radicalisation in local communities, in particular among hard-to-reach Kashmiri communities living in the UK. The project also investigated how Kashmiri youth engage with preventive measures such as the Prevent programme, and how in turn it affects community cohesion in their localities. By conducting interviews with members of the local community and Kashmiri youth, the study applied a bottom-up approach to understand the factors involved in radicalisation. In addition, the study examined the way current prevent policies operate, and formulated some new theories on the Prevent programme's design and delivery processes and how it impacts on national security in the UK (Raiput, n.d.).

22.1. Brief description

Qualitative data analysis methods are used to examine information collected through interviews, focus groups and other qualitative data collection methods. Researchers use data analysis methods to classify and interpret qualitative information.

There are various approaches to qualitative data analysis, including substantive approaches (e.g. grounded theory or thematic analysis, which deal with the content of a narrative) and structural approaches (e.g. conversation or discourse analyses, which are concerned with the structure of a narrative or an account). Another distinction is made

depending on how data is analysed: whether each participant/interview is analysed separately from others (non-cross-sectional analysis) or all interviews are analysed together, providing common sets of topics, codes and themes that are identified and compared across the whole data set (cross-sectional analysis).

Grounded theory is a method that operates 'backwards' compared to traditional research (Strauss and Corbin, 1997): instead of forming a deductive hypothesis before analysis and testing it against collected data, it modifies its hypotheses during the research process. For instance, a researcher examining formulation of violent group, when applying grounded theory, will not use any theories before collecting the data. Instead, he/she will formulate understanding of the group and theories on how the group operates at the same time as conducting analysis. Grounded theory uses systematic comparison of text segments to build a thematic structure and theory from a body of text.

Thematic analysis is a method for identifying, analysing and reporting patterns or themes within data. It is not linked to any pre-existing theoretical framework, and as such it can be used within different frameworks. Thematic analysis organises and describes datasets in rich detail and often interprets various aspects of the research topic (Braun and Clarke, 2006).

There are a range of computer programmes available to help organise segments of qualitative data by theme (e.g. violent groups, P/CVE programmes for teenagers), respondent type (e.g. police officer, community group leader), time and event (e.g. specific P/CVE-related initiative). Qualitative data analysis software, such as NVivo or ATLAS.ti, can be used during the grounded theory coding process to categorise interview transcripts and draw out key findings.

22.2. Purpose

Qualitative data analysis can be conducted for several purposes, including describing a phenomenon in greater detail (e.g. detailed description and analysis of how a violent group operates), explaining similarities and/or differences between cases (e.g. between two or more violent groups), and developing a theory of the studied phenomenon from analysis of empirical material (e.g. a theory of how violent groups are formed, how they select their leaders, recruit new members, etc.). Qualitative data analysis methods are often used either to reduce large data sets to core elements (for instance, only providing analysis of how violent groups finance their activities), or to expand small pieces of data by adding extensive interpretations (e.g. detailed analysis and interpretations of interactions and relationships between violent group members).

Qualitative data analysis methods enable researchers to organise and analyse unstructured qualitative data more systematically to test theories, identify trends and cross-examine information.

22.3. When to use it

The analysis of qualitative data is one in a series of steps in the research process. In some cases, data analysis only starts when all data have been collected and prepared. In other cases, such as grounded theory, the analysis begins at the same time as the data collection and both steps are applied in parallel. For instance, a researcher will start analysing interviews soon after conducting them, rather than only starting analysis when all planned interviews have been conducted.

The next two sections discuss the advantages and limitations of using grounded theory as a method of qualitative data analysis.

22.3.1. Why?

There are several advantages to the use of this approach:

- **Flexibility:** the lack of links to pre-existing frameworks allows thematic analysis to be versatile and adaptive to many purposes. The flexible conceptual framework of grounded theory means that it is relevant to a wide range of field research settings and research methodologies.
- **Creativity:** grounded theory does not start by testing an existing hypothesis but instead uses empirical data to generate original concepts and theories through a creative, inductive process. Thematic analysis can generate unanticipated insights.
- **Systematic analysis:** computer software allows users to classify and organise large quantities of unstructured information, and unpack and examine complex relationships in the data.
- **Analytical breadth/depth:** as an exploratory method, qualitative data analysis is particularly well suited to investigating processes that have attracted little prior attention and where previous research is lacking in breadth and/or depth. This is because it allows researchers to formulate theories based on the collected data rather than testing existing theories.
- **Broad applicability:** qualitative methods are not limited to a specific field, discipline or type of data, and they allow for social as well as psychological interpretations of data. Qualitative data analysis (and thematic analysis in particular) is a relatively quick and easy method to learn and conduct, meaning that it is accessible to researchers with little or no experience of qualitative research.

22.3.2. Why not?

This approach also has various weaknesses, including the following:

- **Managing large volumes of data:** qualitative data analysis tends to produce large amounts of data that can be difficult to manage (for instance, an inexperienced researcher may find it difficult to analyse a large number of interviews). Data reduction/prioritisation can be difficult.
- **Time:** open coding (explained in more detail below) can be a long process. It may require researchers to go back to the same sources of data several times. For instance, there may be several aspects mentioned during an interview that would need to be coded separately, thus requiring multiple rounds of analysis and coding of the same interview.
- **Questionable data validity:** the subjectivity of the data leads to difficulties in establishing reliability and validity of information. For example, interviewees' accounts of the same event may be completely different, making it difficult for a researcher to produce a coherent narrative, but those differences are in themselves findings.
- **Researcher bias:** in qualitative data analysis, there is a risk that the researcher's interpretation of the data may be influenced by their own life experience and prior knowledge of the issue under analysis. For instance, a police officer may interpret data differently than an independent researcher.
- **Limited generalisability of results:** the main goal of qualitative research is to provide a contextualised understanding of the issue under analysis – the ability to draw generalised conclusions from particular instances is considered less important.

For instance, it is more important to understand how a particular violent group recruits new members than to understand general trends in recruitment among a range of violent groups.

22.4. Step-by-step application

This section provides a step-by-step guide on how to analyse qualitative data. It should be noted that in the grounded theory steps 1–3 run in parallel, whereas in thematic analysis these steps are sequential.

1. Conduct data collection and analysis simultaneously
 - Select one or more methods of data collection which are relevant to the research purpose. Data collection methods include (but are not limited to): interviews, focus groups and questionnaires.
 - Gather and analyse data using the selected method(s). For instance, conduct focus groups with members of a local community, and start analysing the data after each focus group has been conducted.
2. Code data through a three-step process
 1. Open coding (comparing incidents applicable to each category):
 - Fragment the data into categories and allocate each issue being examined a code (label). For instance, when analysing violent group recruitment strategies, one code could be recruitment strategies at schools, another could be recruitment strategies in the local community, a third could be recruitment strategies in religious venues, and so on.
 - Code data manually (for instance by colour-coding relevant parts of the text) or using appropriate computer software.
 2. Axial coding (integrating categories and their properties):
 - Re-assemble the data that was fragmented during the 'open coding' process (e.g. put together parts of the text that refer to the same code).
 - Relate subcategories and linked categories, and amalgamate them into a smaller number of overarching categories that explain the data.
 - When reassembling the disaggregated data into broader categories, it is important to consider four factors: (1) the category; (2) the context in which it arises; (3) the actions and interactions that stem from it; and (4) its consequences.
 3. Selective coding (delimiting the theory):
 - After linked categories and sub-categories in the data – in effect re-assembling the raw data – subject the data to 'selective coding' to integrate the data around a central category that 'emerges' from the data (e.g. put together all codes that relate to violent group recruitment strategies).
3. Write memos throughout the research process
 - Through gathering, coding and analysing data, new questions and thoughts about codes and their relationships will arise (e.g. new trends may be observed in how violent groups are organised).
 - Write memos throughout the research process to capture these questions and ideas in order to make codes and categories manageable and to facilitate interpretation of data.

- Memos can include: (i) working definitions of codes or categories; (ii) comparisons between data and between codes and categories; (iii) identified gaps in categories; (iv) fresh ideas and newly created concepts; and (v) comparisons with and links to relevant literature.
4. Stop data collection and analysis when the study reaches theoretical saturation
- Stop data collection and analysis when gathering fresh data no longer sparks new insights or reveals new properties of the theory and its categories or concepts. For instance, this may be when interviews no longer provide any new information on strategies used by radicalised groups to recruit new members.
 - To determine whether saturation has been reached, ask: (i) Are there any gaps in the grounded theory or in its categories? (ii) Are there any vague or underdeveloped definitions? (iii) Are there any data missing? (iv) Are the findings coherent?
5. Write up results
- Integrate memos to complete the analysis
 - Write a report summarising study aims, methodology, results and conclusions.

- There is a risk of researcher bias, meaning that the end result of grounded theory might be disputed by respondents.

List of references

Braun, V. and Clarke, V. 2006. 'Using thematic analysis in psychology'. *Qualitative research in psychology*, 3(2):77–101.

Rajput, O. n.d. *Counter-terrorism policy and radicalisation among UK citizens of Kashmiri heritage*. As of 1 September 2017: http://www.academia.edu/3465131/Counter-Terrorism_Policy_and_Radicalisation_Among_UK_Citizens_of_Kashmiri_Heritage

Strauss, A., and Corbin, J. 1997. *Grounded theory in practice*. Thousand Oaks, CA: Sage.

Check Appendix: Additional resources for further reading.

22.5. Ethics

When conducting qualitative data analysis, the following ethical points should be considered:

- Depending on the data collection methods (observational techniques, interviews, focus groups, surveys), the ethical issues that are mentioned in the chapters on those methods are also relevant for qualitative data analysis and grounded theory.



CHAPTER TWENTY-THREE

Quasi-experimental designs

Megan Sim and Anke van Gorp

See also:



23.1. Brief description

Quasi-experimental designs have similar aims to randomised experiments, namely estimating a causal effect, but lack random assignment of units (e.g. individuals, groups) to different conditions (e.g. receiving a P/CVE intervention versus not receiving the intervention).

Quasi-experiments are also known as non-randomised experiments and cover a wide range of approaches.

Quasi-experiments compare the outcomes of those receiving an intervention or treatment to a comparator group that is as similar as possible to the intervention group in terms of their baseline (i.e. pre-intervention) characteristics. Quasi-experimental designs include:

- Those that use a comparison group:
 - Comparison group designs, in which the outcomes of those exposed to two of more conditions are examined but the evaluator does not control assignment to conditions;
 - Matched group designs, in which the researcher uses information about the units (e.g. demographic variables, psychological traits) to ensure that the group receiving treatment and the group not receiving treatment are as similar as possible (using methods such as propensity score matching), in order to mimic an experimental design;
 - Regression discontinuity designs, in which the evaluator uses a cut-off score on a measured variable to determine eligibility for treatment, and then compares the outcomes of the group receiving treatment to the outcomes for the group not receiving treatment.
- Those that do not use comparison groups:
 - Single-group interrupted time-series designs, in which an effect or outcome is recorded in many consecutive observations over time (e.g. 100 observations), and an intervention is introduced in the midst of those observations to demonstrate its impact on the outcome through a discontinuity

in the trend after treatment. The national implementation of a new immigration policy on the same day is an example of such an intervention. (Note that it is also possible to have a comparison group design and use ITS.)

- Single-case designs, in which an individual is repeatedly observed over time (usually on fewer occasions than in an interrupted time-series design), with the treatment manipulated to demonstrate that treatment affects the outcome.

23.2. Purpose

The objective of quasi-experimental designs is to make causal inferences, meaning that conclusions are drawn about the effect or outcome of specific treatments. For instance, an evaluation about a P/CVE programme could compare the outcomes of a group that received a P/CVE intervention with the outcomes of a group that did not.

There are three requirements for making a causal inference from a quasi-experiment which apply to all causal relationships:

1. The cause must precede the effect. In the case of an evaluation of a P/CVE intervention, the intervention must take place before the measured effect (e.g. number of police contacts in a six-month period following the intervention).
2. The cause must covary with the effect. For instance, if it is expected that longer P/CVE interventions are more effective than shorter interventions, the length of the treatment should be positively associated with the effectiveness of the intervention.
3. Alternative explanations for the causal relationship must be ruled out. Experiments use random assignment of units (e.g. individuals) to conditions (e.g. receiving a

P/CVE intervention or not) to ensure that alternative explanations are distributed over the different conditions and are therefore unlikely to account for the causal relationship. Because quasi-experimental designs do not use random assignment, they rely on other principles to address this requirement, including identification and study of plausible threats to internal validity, primacy of control by design and coherent pattern matching (Shadish et al. 2002).

23.3. When to use it

Quasi-experimental designs can be used prospectively or retrospectively. Quasi-experimental designs can also be used while a P/CVE intervention or programme is ongoing, particularly in cases where the intervention takes place over a long period of time and where it is useful to have preliminary analyses and results. However, it is always recommended that evaluation planning takes place before an intervention or programme of interventions. This is particularly important in the case of a quasi-experimental design, as baseline data should be collected from participants before the intervention or programme takes place.

23.3.1. Why?

The advantages of quasi-experimental designs include the following:

- They can be used to estimate causal effects in cases where it is not possible to randomise individuals or groups into intervention/treatment and control groups. This may be due to ethical concerns, for instance, for withholding or delaying potential effective interventions from a group or individuals (although there are experimental designs that can deal with such issues and there are ethical arguments for using experiments).

- They are particularly useful in circumstances where people choose or are chosen to participate in a programme. However, such techniques rest on the understanding that the reasons for choices are well understood, which may not always be the case (Shadish, 2013).
- They are useful when a policy has been implemented at scale, and administrative data can be used to assess outcomes.

23.3.2. Why not?

The following are some of the drawbacks of using quasi-experimental designs:

- The lack of high-quality data may impede the drawing of causal conclusions. A quasi-experimental design that is implemented after a P/CVE intervention or programme has started may be limited by the lack of baseline data.
- A comparison group that has been poorly matched to the intervention group will affect the quality of the findings. The comparison group needs to be as similar as possible to the intervention group before the intervention (at baseline).

23.4. Step-by-step application

1. Determine the outcome(s) of interest to the evaluation
A non-exhaustive list of outcomes of a P/CVE intervention may include the effectiveness of the intervention, the number of police contacts an individual has, and the beliefs and behaviours of individuals.
2. Decide on the quasi-experimental design that is appropriate for the evaluation
Potential designs include non-equivalent control group designs, interrupted time-series designs, regression discontinuity

designs, single-case designs and matched comparison groups.

Deciding on the quasi-experimental design will also help to identify the control group that should be used in the evaluation. For instance, in an interrupted time-series design, the control group will be made up of multiple observations of the intervention group prior to the intervention. For a matched comparison group, the comparator cases could be drawn from a larger pool of individuals for whom data are available but who did not receive the intervention.

In comparison group designs, the comparator group can be constructed by matching individuals in the intervention group to individuals not undergoing the intervention based on observed characteristics (e.g. age, gender, religion, socio-economic status) using statistical techniques.

3. Collect baseline data

Baseline data should be collected from the intervention and comparator groups on the outcomes of interest prior to the P/CVE intervention. For example, attitudes towards political violence. In some instances this can be collected retrospectively, for example if using administrative records on arrests, but attitudinal data should not be collected retrospectively.

4. Carry out intervention/treatment

5. Collect post-intervention data

Data on the outcomes of interest should be collected from the intervention and comparator groups after the intervention has taken place.

Depending on the quasi-experimental design, outcomes data may also be collected while the intervention is taking

place. That is, the outcome data could be collected at the end of intervention, or several weeks/months post-intervention completion.

6. Analyse the data

Single difference impact estimates compare the outcomes in the intervention group with the outcomes in the comparison group at a single time point following the P/CVE intervention. The difference-in-differences method compares the changes in outcomes over time between the intervention and comparison groups to estimate impact.

Using regression-based methods, single and double difference impact estimates may be made through ordinary least squares regression. For instance, propensity score matching (PSM) approaches look at the difference in means between intervention and comparison groups once the groups are matched. Once intervention and comparison groups have been matched using PSM, it is possible to conduct statistical tests such as regressions and difference tests (e.g. t-tests).

23.5. Ethics

Quasi-experimental methods offer practical options for conducting impact evaluations in real-world settings. By using pre-existing or self-selected groups such as individuals who are already participating in a programme, these methods avoid the ethical concerns that are associated with random assignment – for example, the withholding or delaying of a potentially effective treatment or the provision of a less effective treatment for one group of study participants.

Another ethical concern is related to consent. In the context of P/CVE, consent to treatment

may be difficult to obtain. If an intervention is going to happen anyway (e.g. as part of a court order) and it is a case of deciding between two approaches, then this issue may be negated. However, if a participant has to follow an intervention this does not mean that this participant has automatically consented to participating in the evaluation of this intervention. It should be possible for participants to follow an intervention without participating in the evaluation of the intervention. Using anonymous data would minimise this problem, but also under the GDPR there may be an appropriate legal basis for the data-sharing to take place that does not require consent.

For non-compulsory treatment, it is important to obtain consent for participation in the evaluation in a non-coercive manner and participants should be able to withdraw from the intervention and/or the evaluation without being penalised for doing so.

Depending on the methods used to collect data (e.g. surveys, interviews, focus groups) other ethical issues might be relevant. These are outlined in the chapters on those methods.

List of references

Shadish, W.R., Cook, T.D. and Campbell, D.T. 2002. *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.

Shadish, W. R. 2013. 'Propensity score analysis: promise, reality and irrational exuberance'. *Journal of Experimental Criminology*, 9:129–144.

Check Appendix: Additional resources for further reading.



CHAPTER TWENTY-FOUR

Randomised control trials

Megan Sim, Alex Sutherland and Anke van Gorp

See also:



Example(s) of use

In a study which investigated the role of beliefs about the acceptability of aggression against Jews participants were assigned to either a short educational intervention, aiming to improve intergroup relations, or to a control group (Amjad and Wood, 2009). To determine who would join an extremist group participants were asked to fill in self-report attitude questionnaire pre and post intervention. Those in the intervention group were much less likely to agree to join the extremist group compared to the control group.

24.1. Brief description

Randomised control trials (RCTs) are a comparison group design that assigns units (people, neighbourhoods, schools) to different treatment conditions by chance, akin to flipping a coin to decide who should receive an intervention or not. RCTs are based on the idea of asking, 'What would have happened otherwise?' This is also known as a counter-factual approach. The counter-factual approach underpins both experimental and quasi-experimental studies.

In a simple 'two-arm' trial there are two conditions, 'treatment' and 'control', where the

control group typically does not receive the intervention. It is possible to have so-called 'multi-arm' trials with several conditions being compared at once. For example, there might be several versions of a particular P/CVE intervention that have different components, so the focus is on examining the effectiveness of 'P/CVE versus no P/CVE', but also 'P/CVE option A versus P/CVE option B'.

The benefit of randomisation is that the two groups will be, on average, equivalent in every way apart from whether they were offered the intervention. As a result, randomisation makes it possible to be confident that any differences are the result of the intervention rather than

another factor. There are many different ways in which RCTs can be conducted, but at the heart of any RCT is a randomisation process. Without such a process, a study is not an RCT. Whatever randomisation process is used should be transparent and clearly described. Below are a few examples of approaches to RCTs that cover the unit of randomisation (i.e. what is being randomised) and possible methods of randomisation:

- Individual randomisation: individual units (e.g. people), are allocated to treatment/control groups. Outcomes are measured at the individual level (e.g. attitudes to violent extremism).
- Cluster-randomisation: a collection of units (e.g. neighbourhoods with people in them) are allocated to groups. The focus is still on the effect at the individual level, so outcomes are again measured at that level (e.g. attitudes to compliance with police investigations among citizens from particular neighbourhoods).
- Wait list design: instead of the control group 'missing out' on treatment, the randomisation is for 'treatment now' or 'treatment later', with 'later' being partly determined by how long effects will take to materialise. This might be particularly useful if a service is over-subscribed or has a limit on capacity.
- Trickle randomisation design: instead of having a large pool of units to randomise from the start, it may be that units are infrequent but must be assigned immediately for ethical or practical reasons (Shadish et al. 2002). For example, those identified as being at high risk of becoming radicalised could be assigned to a new therapeutic approach or 'treatment as usual', with the outcomes of the two groups being compared. Such a design may be more appropriate for P/CVE

interventions where participants may become eligible for an intervention in small numbers over many months.

- Stratified randomisation: individuals are divided according to a particular characteristic. For example, individuals may be divided based on risk of radicalisation) and split into 'low', 'medium' and 'high' groups. Randomisation then takes place within each group. This ensures a good balance of risk levels across treatment and control groups.

24.2. Purpose

The objective of experimental designs is to make valid causal inferences; that is, to draw robust conclusions about the effect of an intervention. For instance, an evaluation of a P/CVE programme could compare the outcomes of a group that received a P/CVE intervention with the outcomes of a group that did not. There are three requirements for making a causal inference from a RCT which apply for all causal relationships:

1. The cause must precede the effect, meaning that the intervention must take place before the measured effect (in the case of an evaluation of a P/CVE intervention, the effect may be the number of police contacts in a six-month period following the intervention). The randomisation process ensures that the cause comes before the effect.
2. The cause must covary with the effect, or rather the presence of an intervention should coincide with changes in outcomes. For instance, if it is expected that longer P/CVE interventions are more effective than shorter interventions, the length of the treatment should be positively associated with the effectiveness of the intervention. Covariation will depend on whether or not

the intervention is effective, which is what the experiment is trying to ascertain.

3. Alternative explanations for the causal relationship must be shown to be implausible. Experiments use random assignment of units (e.g. individuals) to conditions (e.g. receiving an intervention or not) to ensure that alternative explanations are distributed over the different conditions and are therefore unlikely to account for the causal relationship. Randomisation ensures that effects are not due to other explanations as 'all else is equal' (on average) between treatment and control groups.

24.3. When to use it

RCTs are used prospectively as it is not possible to randomise after the event. Even if an approach has been in place for some time, randomisation could be introduced (although this may be difficult to implement). However, if there are changes to an approach, system or administrative process, randomisation could be introduced as a way of testing the new approach. This is particularly salient if no-one knows whether or not the changes will be effective. Randomised designs require pre-planning and 'buy-in', and typically involve a baseline or pre-test to assess the equivalence of groups after allocation.

24.3.1. Why?

There are a number of arguments in favour of using RCTs:

- If we do not know whether an approach 'works', experimental designs allow for robust causal inferences to be made. For instance, in the case of an evaluation of a P/CVE programme, experimental designs are able to assess whether or not an intervention achieves its objectives ('Does it work?').
- If there is a mixture of evidence both for and against an approach from different research designs (sometimes called equipoise) then an RCT can contribute to discussions about effectiveness. Note that a single RCT is not conclusive proof of effectiveness, as it depends on the size and quality of a study and whether it can be replicated in other circumstances. Also note that one can argue that there is equipoise if there is no evidence.
- Experimental designs are the fairest way to allocate units if a service is over-subscribed, and appeals to fairness can be one way to make the case for an RCT.
- Sometimes (rarely) it is possible to capitalise on a so-called 'natural experiment' where a random process outside the control of an experimenter generates intervention and control groups.

24.3.2. Why not?

The drawbacks of RCTs include the following:

- A vague or poorly defined outcome and/or intervention makes it very unlikely that an RCT, even if well implemented, would provide useful evidence.
- A lack of buy-in will make implementation difficult, and because the majority of effort goes into ensuring the study is set up in advance, poor implementation means that all the set-up work is wasted. Hence, it is important to ensure that all parties are 'invested' in the RCT design and understand it.
- If it is not possible to collect data directly from those participating in an RCT (e.g. questionnaires capturing attitudes) then administrative data may not be of sufficient quality or relevance for use. If using administrative data then outcomes measured should be subject to direct

influence by the intervention (based on the theory of change).

- There may be few (e.g. under 50) units to be randomised (in the case of an individual-based approach). With a cluster-randomised design, sample size both at cluster (e.g. neighbourhood) level and within clusters (e.g. survey respondents) are important. An underpowered experiment is not necessarily wasteful if it is at a very early stage of development – the question is whether the results are in the expected (i.e. beneficial) direction. It is better to find this out with a small study first, rather than involving hundreds of people only to find that the intervention makes things worse.

24.4. Step-by-step application

1. Identify the policy or approach to be assessed

This could be treatment as usual versus a new intervention, or a comparison of two approaches.

2. Determine the outcome(s) of interest to the evaluation

A non-exhaustive list of outcomes includes the effectiveness of a P/CVE intervention relating to: the number of police contacts an individual has, their beliefs, their behaviours, and contact with particular groups. These outcomes should be based on a sound theory of change, or existing evidence.

3. Decide on the experimental design that is appropriate for the evaluation

- Decisions about study design are driven by several factors including, but not

limited to: the nature of the planned intervention, the unit of randomisation, the number of units available, whether there is a danger of 'contamination' between treatment and control groups,⁸ the amount of funding available, and logistical or resource constraints.

- One of the most important factors in planning an RCT is the anticipated (or desired) effect size (e.g. the difference between treatment and control groups in terms of their likelihood of arrest). A difference of five percentage points (e.g. 20% vs 25%) would require a larger sample than a difference of fifteen percentage points (20% vs 35%).
- Assuming an RCT is appropriate, potential designs include individual or cluster-randomised approaches and wait-list designs. Wait-list designs can mimic how policies are sometimes rolled out, but randomising the order in which areas start implementation.
- It is important to undertake a power calculation (i.e. work out how many units are needed to detect a particular effect size) prior to planning a RCT. Alternatively, this can be a calculation, within the constraints of the study, of the minimum detectable effect size (MDES) that the study could find. All the factors above, including study design, effect size and unit of randomisation, play a part in power calculations.

4. Register trial and write trial protocol

After the decision to pursue a RCT, best practice is to register it in a trials database, and to write and (ideally) publish a trial

8 The risk of 'contamination' depends on whether the groups could interact and the intervention could 'spread' from the treatment group to the control group. In some instances, this 'spread' of an intervention via a network could be desirable, but it would need to be a planned part of an intervention (e.g. choosing 'key players' to 'turn' or influence then assessing the extent to which deradicalisation messages spread through a network).

protocol.⁹ The protocol pre-specifies the outcome(s), intervention, study design and planned analyses (including subgroup analyses). Pre-specifying these protects the integrity of the experiment through reducing the likelihood of repeatedly analysing data until a 'significant' result is found (also known as 'fishing expeditions').

5. Collect baseline data

Baseline data on the outcomes of interest should be collected from the intervention and comparator groups prior to randomisation.

6. Randomise using a pre-specified and well-considered approach

Whatever approach is taken to randomisation should be pre-specified and suitable for the intervention and overall study design. The simplest randomisation is a coin flip, but this is not particularly scientific because it is possible to manipulate coin tosses and this has largely been discredited as an approach (Clark and Westerberg, 2009). It is possible to use widely available software such as Microsoft Excel to randomise, but there are also online tools that can also be used for this purpose.¹⁰ Note that confidential material should not be uploaded to an online randomisation tool.

Best practice for randomisation is to have it carried out by a person or organisation that is demonstrably independent of the intervention or delivery team (e.g. not colleagues or relatives). This is to allay suspicion of someone consciously or

unconsciously biasing the randomisation. Having a third party involved also strengthens the credibility of the evaluation.

7. Assess success of randomisation

Compare treatment and control groups on baseline data in terms of their means and distribution of measures. Note that this does not mean statistically testing differences – any differences between groups should be the result of chance if the randomisation was successful.¹¹ If there are lots of differences between groups (e.g. the average age of the treatment group is 25 and for the control group it is 55, and most of one group are males) it might be that the randomisation has not worked as it should. In that case, it is important to ask an external expert to review what has been done prior to starting the intervention (and, if necessary, re-randomising before the intervention begins).

8. Monitoring intervention delivery

This phase could include a process evaluation to understand more about the actual delivery of the intervention. Depending on how extensive or complex the intervention is, the scale of the process evaluation may vary. Simplistic approaches such as text messages do not need a process evaluation element, but P/CVE interventions are unlikely to be so 'light touch'. As such, it is important to understand the level of compliance with the intervention plan (e.g. if six appointments should have been offered by providers, were they?) and the fidelity of the intervention

9 E.g. ISRCTN trial registry (<https://www.isrctn.com/>), or sending protocols to a journal such as Trials (<https://trialsjournal.biomedcentral.com/>) or uploading the protocol to a repository such as Crim-PORT (<http://www.crim.cam.ac.uk/research/experiments/rex-post/crimport.pdf>).

10 E.g. <https://www.random.org/lists/>

11 See <http://www.consort-statement.org/checklists/view/32--consort-2010/510-baseline-data>

(e.g. if the appointments have pre-specified plans, were these followed as intended?).

9. Collect post-intervention outcome data

Data on the outcomes of interest should be collected from the intervention and comparator groups during and after the intervention has taken place. In practice, trials usually involve waiting for a pre-defined period after the intervention has been completed to collect data, but this will vary depending on the budget and logistics of data collection. Collecting data from administrative sources can take longer, and there are special considerations for some types of data (e.g. reconviction data) (see for example Sutherland, 2013).

10. Analyse the data

Compare the treatment and control groups across the outcome(s) of interest. Analyses should take account of the study design used, so if a cluster-randomised design was used then this should be taken into account. One way of analysing RCTs is an 'intention-to-treat' (ITT) approach in which everyone is analysed according to their allocation, regardless of whether they complied with the intervention or not.

24.5. Ethics

A common ethical objection to RCTs is that they involve withholding intervention or giving someone a sub-optimal intervention. Two responses to this are that (a): comparisons would normally be made to 'treatment as usual' versus 'new treatment; and (b) it is not known if the intervention is beneficial.

Another problem is that the reason for conducting an RCT is that it is not known whether an intervention will work. It might be considered unethical to test interventions with no knowledge of their effectiveness. However, to actively test an approach that is known to cause harm is ethically problematic (e.g. the 'scared straight' programme is known to cause harm, but is sometimes revisited as a 'new idea' by new policy-makers).¹²

Another ethical concern is related to consent. In the context of P/CVE, consent to treatment may be difficult to obtain. If an intervention is going to happen anyway (e.g. as part of a court order) and it is a case of deciding between two approaches, then this issue may be negated. However, if a participant has to follow an intervention, this does not mean that this participant has automatically consented to participating in the evaluation of this intervention. It should be possible for participants to follow an intervention without participating in the evaluation of the intervention, without services being denied.

For non-compulsory treatment, it is important to obtain consent for participation in a non-coercive manner and participants should be able to withdraw from the intervention and/or the evaluation without being penalised for doing so.

Depending on the methods used to collect data (surveys, interviews, focus groups) other ethical issues might be relevant. For details, see the chapters covering those methods.

12

See for example: http://www.cochrane.org/CD002796/BEHAV_scared-straight-and-other-juvenile-awareness-programs-for-preventing-juvenile-delinquency.

List of references

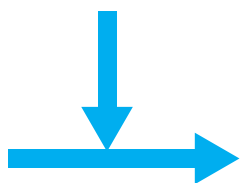
Amjad, N. and Wood, A.M., 2009. 'Identifying and changing the normative beliefs about aggression which lead young Muslim adults to join extremist anti-Semitic groups in Pakistan'. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression*, 35(6), pp.514-519.

Shadish, W.R., Cook, T.D. and Campbell, D.T. 2002. *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.

Clark and Westerberg. 2009. 'How random is the toss of coin?'. *CMAJ*, 181(12): E306–E308.

Sutherland, A. 2013. *A Methodology for Reconviction using Police National Computer (PNC) data*. Unpublished working paper.

Check Appendix: Additional resources for further reading.



CHAPTER TWENTY-FIVE

Realist evaluation

Megan Sim and Anke van Gorp

See also:



25.1. Brief description

Realist evaluation is a theory-driven approach that can be used to evaluate programmes and interventions (other theory-based approaches include contribution analysis, theory of change and policy scientific approach). Realist evaluation considers that programmes and interventions are based on an underlying programme theory, which describes how, under which conditions, and for whom the intervention is expected to lead to its desired outcomes. Therefore, realist evaluation can be used to understand how programmes work by testing their theory and considering whether this theory is thorough, plausible, durable, practical and valid.

Realist evaluation highlights four linked concepts for understanding how programmes work: 'mechanism', 'context', 'outcome pattern' and 'context–mechanism–outcome pattern configuration'. These are explained below.

Mechanism refers to what it is about interventions that bring about any effects. Examining mechanisms involves assessing the process of how stakeholders and participants

interpret and act upon the intervention strategy. For instance, an intervention targeted at potentially radicalised youth that comprises group sessions has several potential mechanisms. The intervention may increase the youths' abilities to think critically by offering an outlet for open and non-judgemental debate. It may improve participants' social skills by creating a space for interacting with people from different backgrounds. On the other hand, it may encourage further radicalisation by fostering a hostile dynamic between the group facilitators and the participants. A realist evaluation would consider different mechanisms to trace the logic of an intervention.

Context refers to conditions in which an intervention is introduced that are relevant to the programme mechanism. For instance, contexts relevant to the group intervention introduced above may include participants' readiness to accept anti-radical ways of thinking, the extent of their ongoing contact with radicalising factors, recent policy developments, and the support of their friends and families. A realist evaluation would aim

to distinguish contexts that support the programme theory from contexts that do not.

Outcome patterns refer to the intended and unintended consequences of programmes as a result of different mechanisms acting in different contexts. Programmes may result in a variety of outcome patterns; when programmes are implemented, the mechanisms activated may vary depending on different conditions, resulting in multiple contexts. A realist evaluation would consider the intended and unintended consequences to disentangle how the programme has worked.

Context–mechanism–outcome pattern configuration (CMOC) brings together the assessment of different mechanisms and contexts to predict and explain the variety of outcome patterns. A realist evaluation would construct CMOC models of how the programme works that can be tested empirically.

25.2. Purpose

Realist evaluation is used to understand how programmes work. Rather than solely focusing simply on whether, for instance, an intervention worked or not, realist evaluation focuses on identifying what aspects of the intervention worked, in which circumstances and for whom. The complete realist question is: ‘What works, for whom, in what respects, to what extent, in what contexts, and how?’ (Better Evaluation, n.d.). The underlying purpose of conducting a realist evaluation is to refine the programme theory for ongoing or future iterations of the programme or intervention.

25.3. When to use it

Realist evaluation can be used prospectively (before an intervention is implemented), concurrent to the programme or retrospectively (after the intervention was completed).

25.3.1. Why?

There are a number of arguments in favour of realist evaluation, including the following:

- Realist evaluations are suitable for assessing how interventions work in complex situations because they aim to deconstruct the causal relationships underlying interventions.
 - For example, in the evaluation of a disengagement intervention for people who subscribed to an extremist ideology, a realist evaluation would consider not simply how many people disengaged, but rather why the intervention worked to help people disengage. Questions to consider could include which aspects of the intervention were conducive for helping which groups of people to disengage, and how effective the disengagement process was.
- Realist evaluations provide findings that could be useful to practitioners and policymakers because they aim to understand how interventions work and the conditions that are necessary for an intervention to work.
 - For instance, a realist evaluation could be useful in understanding that a school-based intervention was able to change the attitudes and behaviours of radicalised youths, but that this intervention would only work in urbanised areas and would have to be modified for use in other contexts.

25.3.2. Why not?

Arguments against the use of realist evaluation include the following:

- While a realist approach to evaluation may not be more resource- or time-intensive than other theory-based evaluations, it

can be more expensive than a pre/post evaluation design.

- When conducting a realist evaluation, sufficient time and resources need to be dedicated to assessing the interactions between the intervention, the different people involved (e.g. intervention managers and frontline staff, participants, participants' friends and family) and the context.

25.4. Step-by-step application

1. Develop the initial programme theory/theories

The first step in conducting a realist evaluation is to develop the programme theory or theories to be tested. The programme theory describes how and under which conditions the intervention will result in its desired effects. In order to develop the initial programme theory, the evaluators can examine, for example, public or internal documentation, stakeholders who have designed the intervention, practitioners and intervention staff, as well as previous evaluations and academic literature. The evaluators should explicitly identify the expected mechanisms, contexts, outcome patterns and CMOCs at this stage so that data collection can focus on testing the different elements of the programme theory. This is similar to constructing a ToC model that links the context of an intervention to activities and results in order to explain how and why the desired change is expected to happen.

2. Collect the data

The next step in a realist evaluation is to collect the data that will enable the initial programme theory to be tested. Realist evaluations typically make use of both quantitative and qualitative data, with the

former being focused on contexts and outcomes and the latter being focused on mechanisms. Quantitative data that could be collected may include administrative records of intervention participants (such as police or school data); qualitative data may include interview and focus group data from relevant stakeholders (e.g. intervention staff, participants, participants' friends and family, the wider public).

3. Conduct a realist data analysis

The third step in a realist evaluation is to test the hypothesised mechanisms, contexts, outcome patterns and CMOCs using the data collected. In a realist evaluation, it is expected that there will be a nuanced outcome pattern of successes and failures across the programme. The purpose of the analysis is to assess if the outcome patterns can be explained by the hypothesised CMOC models.

4. Assess and interpret the analysis

The final stage is to consider whether the programme theory is supported or refuted by the analysis. It should be noted that findings are unlikely to be unequivocal – while some outcome patterns may be attributable to certain contexts and mechanisms, others may be unclear. Unanticipated outcomes may require the evaluators to re-consider the initially developed programme theory. Therefore, this final stage is an iterative process and further rounds of analysis may be necessary.

25.5. Ethics

When conducting a realist evaluation, the following ethical points should be considered:

- Ethical issues may arise depending on the methods used to gather data. See for example chapters on interviews, surveys,

focus groups, literature review and case studies.

- A potential ethical issue specific to theory-based evaluation is the difficulty of reconstructing the programme theory without falling prey to hindsight bias (Roese and Vohs, 2012). Stakeholders might be affected by what they have experienced and learned during the intervention and therefore might not be able to correctly reconstruct assumptions and ideas they had during the design the intervention under evaluation. Moreover, if they have invested in a project already, the 'sunk cost fallacy' might also blind stakeholders (Dobelli, 2013).

List of references

Better Evaluation n.d. 'Realist evaluation'. As of 18 February 2018:

http://betterevaluation.org/approach/realist_evaluation

Dobelli, R. 2013. *The art of thinking clearly*. London: Sceptre.

Roese, N. J. and Vohs, K.D. 2012. 'Hindsight Bias'. *Perspectives on Psychological Science*, 7(5):411–426.

Check Appendix: Additional resources for further reading.



See also:



CHAPTER TWENTY-SIX

Stakeholder analysis

Joanna Hofman, Kei Ito and Anke van Gorp

Example(s) of use

Options UK, on behalf of the London Borough of Waltham Forest, performed a stakeholder mapping documented in their 2010 report 'Understanding East London's Somali Communities'. The research which formed part of a wider East London Alliance programme, 'Building Somali Resilience and Leadership', aimed at helping communities build resilience to violent extremism. Stakeholder mapping was used to identify which key stakeholders to engage in the research process (Options UK, 2010).

In its report 'Strengthening Capacity to Prevent Violent Extremism in the Kyrgyz Republic', Search for Common Ground employed stakeholder analysis to examine the relationship between state authorities and community leaders. The stakeholder analysis was also used to understand each stakeholder's position and interest in relation to four indicators pertinent to the objectives of the study in question (power relations, values, relationships in the intra-Islamic sphere and platform for dialogue) (Search for Common Ground, 2013).

26.1. Brief description

Stakeholder analysis (sometimes referred to as stakeholder mapping) is an analytical tool to map and understand the power, position and perspectives of the actors who have an interest in, and/or are likely to be affected by, a particular policy, programme or piece of legislation. Stakeholder analysis helps to

systematically identify and group primary stakeholders (i.e. those who are directly affected by a particular intervention, e.g. beneficiaries) and secondary stakeholders (who are only indirectly affected but can play a role in implementation arrangements). For instance, in a P/CVE intervention, key stakeholders may typically include programme clients and/or programme facilitators, the

families of clients, governmental agencies, community-based organisations including schools, religious and local initiatives, as well as sponsoring organisations and others. These actors can be primary or secondary stakeholders depending on the nature of the intervention.

26.2. Purpose

The purpose of a stakeholder analysis is to indicate whose interests, views and opinions need to be considered, as well as why and how those interests should be taken into account. Given that stakeholders involved in a particular intervention might have provided specific expertise, resources or inputs, it is necessary to understand what incentives they had for getting involved, whether they were likely to benefit (or not) from the initiative, what role they played and what their perceptions of the intervention are. In the P/CVE context, a stakeholder analysis is a particularly useful tool as these types of interventions commonly involve and affect a great variety of actors at different levels, as outlined above.

In terms of planning a P/CVE policy or programme, stakeholder analysis can also be used to reveal power imbalances among weaker stakeholder groups¹³ and (if possible) reduce power imbalances in the design and/or implementation phase. In addition, stakeholder analysis is a useful instrument to manage and engage with stakeholders.

26.3. When to use it

Stakeholder analysis can be used in various ways. It can be used retrospectively as a tool to assess how different groups of stakeholders were affected by an intervention.

In terms of data collection, it can be used as a mapping instrument in order to help determine which data should be collected from which stakeholder and at which point in the evaluation process.

Stakeholder analysis may also be applied in project planning and policy development as it may help to develop a better understanding of stakeholder needs and thus enable better-designed P/CVE interventions. Lastly, stakeholder analysis can be employed on a continuous basis to understand how actors and relationships change before, during and after the implementation of the intervention.

26.3.1. Why?

Stakeholder analysis has the following advantages:

- It helps evaluators to get to know stakeholders better and manage relationships (e.g. relationships between local community leaders and local government).
- It may help identify potential risks and challenges and inform a decision-making process.
- It may facilitate engagement with stakeholders and their acceptance of (planned) interventions and evaluation outcomes.

26.3.2. Why not?

Disadvantages of stakeholder analysis include the following:

- Results may become obsolete if too much time passes before they are reviewed

13

For example, who is involved in discussions, who is not and why?

- Assessment of analysis may be subjective and it may not (equally) cover all groups of stakeholders.

26.4. Step-by-step application

The most popular type of stakeholder analysis focuses on two dimensions: power and interest. More sophisticated techniques include a third dimension (attitude) or use dedicated software to help to manage and visualise the stakeholders. Stakeholder analysis can be conducted at different points in time (prior to, during, and after an intervention being implemented) in order to either facilitate the design and implementation of a particular intervention or to assess its effectiveness.

Stakeholder analysis usually involves four main steps:

1. List all possible stakeholders and identify their key characteristics (e.g. internal or external to the intervention, non-governmental or governmental, etc.).
2. Consider the (positive or negative) impact of the intervention on the identified stakeholders.
3. Identify each stakeholder's interest (e.g. financial or emotional) and attitude (e.g. engaged or disengaged) in relation to the intervention.
4. Decide which stakeholder groups should participate, in what form, at what level and at which point during the evaluation.

26.5. Ethical considerations

Given that stakeholder analysis may determine which groups of actors will participate in and which will be excluded from an evaluation, ethical considerations are particularly relevant when using this tool.

Informed consent is a crucial ethical concern in relation to this method. In a stakeholder analysis, information about non-participants can be gathered if participants who are linked to them are involved in the research. Even if individuals do not want to participate in the research and provide information, they will still be in the stakeholder analysis with information related to them based on other stakeholders' input or documents. This makes informed consent problematic as there is no real possibility of refusing to provide data, because other stakeholders can provide data about non-participants. Only if a large part of the research population also refuses to participate would an individual's refusal to participate lead to no information being collected about him or her. However, if a large part of the stakeholders were to refuse to give data, stakeholder analysis would become impossible. Borgatti and Molina (2003) claim that this is acceptable, but the interests of stakeholders who are not participants in the evaluation should be taken into account.

For details on ethical issues surrounding methods that can be used during stakeholder analysis (such as surveys, questionnaires, literature review, interviews and observations), see the chapters on those methods.

List of references

Borgatti, S.P. and Molina, J.L. 2003. 'Ethical and Strategic Issues in Organizational Social Network Analysis'. *The Journal of Applied Behavioral Science*, 39(3):337–349.

Options UK. 2010. *Understanding East London's Somali Communities: A study conducted for the East London Alliance*. As of 18 February 2018: <http://karin-ha.org.uk/wp-content/uploads/2013/01/Understanding-East-Londons-Somali-Communities.pdf>

Search for Common Ground. 2013. *Baseline Assessment Report in the Framework of*

Strengthening Capacity to Prevent Violent Extremism in the Kyrgyz Republic. As of 18 February 2018:
<https://www.sfcg.org/understanding-violent-extremism/>

Check Appendix: Additional resources for further reading.



CHAPTER TWENTY-SEVEN

Surveys

Kate Cox and Anke van Gorp

See also:



Example(s) of use

In January 2010, the Danish Ministry of Refugee, Integration and Immigration Affairs initiated a survey on de-radicalisation and disengagement. The objectives of the survey were: 1) to map the practical experiences in EU Member States with policies and programmes; and 2) to foster the exchange of good practice among EU Member States in order to address the challenges of extremism and radicalisation.

The questionnaire was sent to all 27 EU Member States. Three key findings were: 1) all but one of the 18 countries that responded have developed, or are developing, strategies to prevent radicalisation and extremism; 2) of the 18 countries, 13 have experienced problems with right-wing extremism, 12 have experienced problems with left-wing extremism, 10 have experienced problems with militant Islamism, and 7 have experienced problems with separatist movements, animal rights groups and/or other groups; and 3) in most countries analysed, right-wing extremism and militant Islamism are considered to be the most serious threats (Danish Ministry of Refugee, Immigration and Integration Affairs, 2010).

The aim of the UK's programme 'Tackling radicalisation in dispersed societies (TaRDs)' was to achieve a deeper understanding of the potential for dissatisfaction among Sutton's residents and the potential for those individuals to be radicalised towards violent extremist viewpoints (Lewis, 2013). The project used a survey approach to capture views on local sources of dissatisfaction across the borough and the associated drivers of dissatisfaction, and to identify those individuals seeking to exploit that dissatisfaction for the purposes of violent extremism.

27.1. Brief description

A survey is a research instrument that gathers standardised information from a selected sample of individuals or organisations (e.g. police officers or residents). Surveys consist of a sample, a method of data collection (e.g. a questionnaire), and a set of questions. Surveys are often – but not always – designed for statistical analysis of responses. A survey may focus on preferences, opinions, behaviour or factual information, depending on its purpose. Surveys can be divided into two broad categories: the self-completed questionnaire and the interview survey. Self-completion questionnaires can be conducted online, by email and by post, while interview-based questionnaires can be carried out by telephone, face-to-face or via video-calling. There are now many companies that offer professional survey services.

27.2. Purpose

The purpose of a survey is to collect qualitative and/or quantitative information on a sample of individuals or organisations. As surveys often collect comparable information on a large number of cases, they can be used to provide baseline data against which the performance of a strategy, programme or project can be compared. Surveys are therefore often used to inform the formal evaluation of the impact of a programme or project. For instance, in the P/CVE context, a range of stakeholders such as trainers delivering a programme, programme participants and wider community members may be asked to fill in a survey questionnaire asking them about their assessment of the programme implementation. Collecting a large number of views and perspectives makes it possible to compare and contrast views of particular stakeholders and identify common benefits and challenges highlighted by respondents. These findings, in turn, can

be used to assess the implementation of a specific P/CVE programme and prepare recommendations on how the implementation could be improved in the future.

When collected from a representative random sample of a population, surveys can be used to generalise findings from a sample to wider population the sample was drawn from. As an example, general population surveys in the UK typically consist of a small number of people (around 1,000) whose views are representative of the general population provided that they are sampled well (i.e. based on sound methodology). In the P/CVE context, a survey might use a random sample of police officers to generalise findings to the wider population of police officers.

27.3. When to use it

Surveys are primarily used during the data collection phase of evaluation. During the data collection process, surveys can be used to gather a standardised set of responses to improve the researcher's understanding of one or more key issues (e.g. teachers' views on youth at risk of radicalisation, collaboration between organisations tackling community safety). The set of questions asked is often informed by a literature review or quantitative analysis conducted at an earlier stage of the research process. After the survey has been conducted, findings can be analysed and incorporated into the reporting process.

It can be particularly beneficial to use this research tool in the following circumstances:

- When the boundaries and characteristics of a topic or subject can be easily determined in advance (all survey types): survey questionnaires need to include a relatively small number of well-defined questions that would be easy to understand for a wide range of respondents.

- When fast turnaround is necessary (telephone, online and email surveys): surveys have the advantage of collecting views and perspectives of a large number of respondents in a short timeframe.
- When budget is limited (online and postal surveys): survey deployment and collection of data is a cost-efficient method compared with other methods of collecting information from a large number of respondents.

27.3.1. Why?

The use of surveys has the following advantages:

- It ensures consistent data collection from respondents as all respondents are asked to provide answers to the same questions. This applies to all survey types.
- It is an efficient way of collecting information from a diverse range of stakeholder groups as it makes it possible to gather information from a large number of individuals and organisations. This applies to all survey types.
- It makes it possible to elicit honest views from practitioners and other stakeholders (e.g. students, teachers, police officers, trainers delivering programmes), as the respondent feel that they are talking to a computer rather than a person. This may be particularly important for sensitive topics such as involvement in criminal activity or having strong views on a particular issue (e.g. support for radicalised individuals or groups). The standard approach to sensitive questions in face-to-face surveys is Computer Assisted Personal Interviewing. This point applies to online and email surveys.
- It is faster than other evidence-gathering methods since a large number of respondents can provide their views and

opinions at the same time, e.g. when filling in an online survey questionnaire. This applies to telephone, online and email surveys.

- It is cost-effective as the deployment and data collection cost is relatively low compared with other methods. However, with postal surveys, there is a cost involved in entering data once it is collected. Some companies offer data entry as a service. If data entry is undertaken by evaluators, then time needs to be set aside to check the quality of data entry and use a data-entry 'mask' such as epidata to reduce errors. This applies to online and postal surveys.
- It is flexible and adaptable as the timing of conducting survey interviews can be arranged to suit the availability of a surveyed individual, e.g. in the evening when respondents are home from work. This applies to telephone and face-to-face surveys.
- It avoids interviewer bias as there is no direct contact between researchers collecting the survey data and the respondent providing answers to questions. In this way, researchers' characteristics (e.g. age, sex, ethnicity) do not influence respondents' answers. However, respondents may still feel inclined to give socially desirable responses, even without an interviewer present. This applies to self-completion questionnaires.

27.3.2. Why not?

The drawbacks of surveys include the following:

- Potential for interviewer bias: for example, some people may be more willing to discuss a sensitive issue with a female interviewer than with a male one, for instance when discussing experiencing violence by a partner or how to raise a

child. Computer-aided interviewing, wherein a tablet device with headphones is handed to the interviewee so he/she can answer sensitive questions without the interviewer having to ask or hear the answers, may help. However, issues with potential biases arising as the result of different data collection approaches should also be considered. For example, if people can only respond online, what segments of the population might be missing? This point applies to interview-based questionnaires.

- Researchers cannot be sure whether the respondent has understood the question being asked as researchers usually cannot provide any further clarification on questions once these are posted online. It is advisable to pilot and test survey questions in advance of use, including so-called cognitive testing. This point applies to online and postal surveys.
- Likelihood of low return rates: for instance when potential respondents do not want to engage with the issue under investigation (e.g. when it is a sensitive topic, or when they feel that they do not have the knowledge required to answer the questions) or when there are practical barriers to participation (e.g. older respondents do not have adequate IT skills to answer online questionnaires, or survey distribution channels failing and potential respondents being unaware of a survey being conducted). This point applies to email and online surveys.
- Surveys are inflexible as it is typically not feasible to change survey questions when respondents start providing answers. This applies to quantitative online surveys.
- Geographical limitations and travel costs can be obstacles for researchers responsible for survey data collection. This applies to face-to-face surveys.

27.4. Step-by-step application

This section offers a step-by-step guide to the use of surveys.

1. Determine purpose of survey
 - Define the overarching aim of the survey (e.g. understanding relationships in families with radicalised individuals).
 - The purpose will influence the choice of questions and respondents (e.g. parents, brothers and sisters, and wider family members of radicalised individuals).
2. Select respondents
 - Select respondents with the survey's purpose in mind.
 - Determine sample size (decide how many respondents will have to answer survey questions).
 - Determine sampling approach if applicable (e.g. purposive, simple random, cluster, stratified).
3. Review survey
 - Review survey to determine:
 - The required structure of the survey (e.g. the flow of subject areas that respondents will be questioned about);
 - The number of sections/questions, bearing in mind the physical limitations of respondents (for example, if it is unlikely that respondents would be able to spend more than 15–20 minutes on answering survey questions);
 - The complexity of the questions, ensuring that each question is focused only on a specific and easy-to-grasp concept;
 - The phrasing of the questions (for example, it is best to avoid specialist jargon when surveying to ensure that questions are easy to understand).

- for respondents – again, piloting is crucial);
- The number of stakeholder groups involved;
- Whether different surveys are required for each of the stakeholder groups: for instance, when surveying teachers and students, researchers may prepare separate versions of the questionnaire for teachers and for students. Some of the questions may be identical in both questionnaires, but there may also be questions that are only included in one version.
- Meet with project sponsors to discuss and agree on:
 - Any adjustments to phrasing of questions (this may be done when the piloting of the survey is conducted);
 - The form and function of the survey (e.g. face-to-face, online, postal);
 - The mechanics of implementing the survey (e.g. agreeing on introductory text outlining the survey, identifying documents for respondents to review before completing the survey, etc.).
- 4. Invite stakeholders to participate in survey
 - Invite stakeholders to take part by email and follow-up phone calls
 - Send participants a copy of the survey or provide a link to an online questionnaire.
 - 4.1. Conduct survey (in-person or phone surveys only)
 - Conduct the survey in person or over the phone (depending on survey type).
 - 4.2) Deploy survey (paper-based or electronic surveys only)
 - Upon receipt of agreement to participate, send the survey to participants for self-completion.

5. Follow up respondents (paper-based or electronic surveys only)

Follow up non-respondents by email or telephone once the survey has been deployed for a short time period to ensure participation from all potential respondents.

6. Collate data

Collate data upon survey completion.

7. Conduct analysis

Using survey data, conduct analysis to extract results and conclusions.

27.5. Ethics

Important ethical issues to address when conducting a survey are:

- Confidentiality: the respondent's right to confidentiality should always be respected and any legal requirements on data protection adhered to. Respondents should be informed about survey data confidentiality before they start answering survey questions.
- Informed consent: respondents should be fully advised on the aims of the survey and the time it will take to finish it, so that they may make an informed judgment about whether they wish to participate. Respondents' consent to participate in the survey must be obtained and recorded. With online surveys, this should be a 'tick box' they can check to confirm they understand the research, or a text that says 'by continuing you confirm that...'
- Harm to participants, whether emotional or physical, must be avoided. The amount of time requested from respondents and the emotional burden should not be excessive. This is especially relevant if the survey includes questions about sensitive issues or painful experiences.

List of references

Danish Ministry of Refugee, Immigration and Integration Affairs. 2010. *The Challenge of Extremism: Examples of Deradicalisation and Disengagement Programmes in the EU*. As of 12 November 2017:
<http://www.youthpolicy.org/library/documents/the-challenge-of-extremism-examples-of-deradicalisation-and-disengagement-programmes-in-the-eu/>

Lewis, C., 2013. 'The Safer Sutton Partnership Service: Blueprint for the Future?'. *The police journal*, 86(1), pp.15-28.

Check Appendix: Additional resources for further reading.

Appendix: Additional resources

Case studies	
Hancock, D. R. and Algozzine, B. 2006. <i>Doing Case Study Research: A Practical Guide for Beginning Researchers</i> . New York: Teachers College Press.	An accessible introductory text on case study research. Offers step-by-step guidance to help researchers shape a project from conceptualisation to completion.
Stake, R. 1995. <i>The Art of Case Study Research</i> , Thousand Oaks, CA: Sage.	A concise guide to case study research.
US General Accounting Office. 1990. <i>Case Study Evaluations</i> . As of 18 February 2018: http://betterevaluation.org/resources/guide/case_study_evaluations_gao	Describes how case study methods can be used. Describes six applications of case study methods, including the purposes and pitfalls of each, and explains similarities and differences among the six. Presents an evaluation perspective on case studies, defines them, and determines their appropriateness in terms of the type of evaluation question posed.
Yin, R. 2003. <i>Case Study Research: Design and Methods</i> (3rd ed.). Thousand Oaks, CA: Sage.	Offers comprehensive coverage of the design and use of the case study method as a research tool.
Comparisons/benchmarking	
Better Evaluation. 2014. 'Standards, evaluative criteria and benchmarks'. Better Evaluation website. As of 18 February 2018: http://betterevaluation.org/evaluation-options/benchmarks_standards	A brief introduction to the use of benchmarking in evaluation.
Ling, T. and Villalba van Dijk, L. 2009. <i>Performance Audit Handbook</i> . Santa Monica, CA: RAND Corporation.	An overview of tools applicable to performance audit and evaluation.

World Bank. 2010. 'Benchmarking: A tool to improve the effectiveness of Monitoring and Evaluation in the policy cycle'. World Bank website. As of 18 February 2018: http://siteresources.worldbank.org/INTLAC/Resources/257803-1269390034020/EnBreve_154_Web_Final2.pdf	An overview note from the World Bank on how to apply benchmarking for monitoring and evaluation purposes.
Contribution analysis	
European Commission, Directorate-General for Regional and Urban Policies. 2013. <i>Evalsed Sourcebook: Method and Techniques</i> . European Commission. As of 18 February 2018: http://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/evaluation_sourcebook.pdf	Sourcebook outlining a range of methods and techniques for application in the evaluation of socio-economic development.
Goldwyn, R. and Chigas, D. 2013. 'Monitoring and evaluating conflict sensitivity: Methodological challenges and practical solutions'. As of 18 February 2018: http://www.smallarmssurvey.org/fileadmin/docs/M-files/CCRV/CCVRI-Monitoring-and-evaluating-conflict-sensitivity-challenges-and-solutions.pdf	Gives practical guidance on how to monitor and evaluate interactions between an intervention and conflict.
Mayne, J. 2008. <i>Contribution Analysis: An approach to exploring cause and effect</i> . ILAC methodological brief. Institutional Learning and Change Initiative. As of 18 February 2018: http://www.betterevaluation.org/en/resources/guides/contribution_analysis/ilac_brief	An accessible overview of contribution analysis, authored by its proponent John Mayne.
Vaessen, J. and Raimondo, E. 2012. 'Making sense of impact: A methodological framework for assessing the impact of prizes'. <i>Evaluation</i> , 18:330–347.	Discusses the intervention logic underlying UNESCO prizes, the corresponding framework for empirical assessment and contribution analysis as a tool for an impact evaluation of a prize.
Wimbush, E., Montague, S. and Mulherin, T. 'Applications of contribution analysis to outcome planning and impact evaluation'. <i>Evaluation</i> , 18:310–329.	Journal article outlining benefits of contribution analysis and providing examples of its use.
Cost-benefit analysis	
Drummond, M.F., Sculpher, M.J., Claxton, K., Stoddart, G.L. and Torrance, G.W. 2015. <i>Methods for the economic evaluation of health care programmes</i> . Oxford: Oxford University Press.	An overview of methods used in health economic evaluation, including calculation of CEA ratios.

Eide, E. n.d. '4.8 Cost-benefit analysis (CBA)'. Oslo University. As of 18 February 2018: http://www.uio.no/studier/emner/jus/jus/JUS5831/h12/undervisningsmateriale/h12-day-5-cba.ppt	Briefly describes the steps required to undertake CBA.
European Commission, Directorate-General for Regional and Urban Policies. 2013. <i>Evalsed Sourcebook: Methods and Techniques</i> . European Commission.	Outlines what CBA can be used for, as well as the steps required to conduct it.
HM Treasury. 2011. <i>The Green Book: Appraisal and Evaluation in Central Government</i> . London: HM Treasury.	High-level guidance on evaluation in government (UK).
Martin, A., Ogilvie, D. and Suhrcke, M. 2014. 'Evaluating causal relationships between urban built environment characteristics and obesity: a methodological review of observational studies'. <i>International Journal of Behavioral Nutrition and Physical Activity</i> , 11:1–15.	A guide to the use of econometric methods for the evaluation of interventions where RCTs may not be feasible.
Cost-effectiveness analysis	
Better Evaluation. 2014. 'Cost Effectiveness Analysis'. Better Evaluation website. As of 18 February 2018: http://betterevaluation.org/evaluation-options/CostEffectivenessAnalysis	A brief introduction to CEA and to the calculation of cost-effectiveness ratios.
Drummond, M.F., Sculpher, M.J., Claxton, K., Stoddart, G.L. and Torrance, G.W. 2015. <i>Methods for the economic evaluation of health care programmes</i> . Oxford: Oxford University Press.	An overview of methods used in health economic evaluation, including calculation of CEA ratios.
European Commission, Directorate-General for Regional and Urban Policies. 2013. <i>Evalsed Sourcebook: Methods and Techniques</i> . European Commission.	Outlines what CBA can be used for, as well as the steps required to conduct it.
HM Treasury. 2011. <i>The Green Book: Appraisal and Evaluation in Central Government</i> . London: HM Treasury.	High-level guidance on evaluation in government (UK).
Martin, A., Ogilvie, D. and Suhrcke, M. 2014. 'Evaluating causal relationships between urban built environment characteristics and obesity: a methodological review of observational studies'. <i>International Journal of Behavioral Nutrition and Physical Activity</i> , 11:1–15.	A guide to the use of econometric methods for the evaluation of interventions where RCTs may not be feasible.

Cross-sectional data analysis	
Bourque, L. B. 2004. 'Cross-sectional design' in Lewis-Beck, M.S., Bryman, A. and Liao, T.F. (eds.) <i>The SAGE Encyclopedia of Social Science Research Methods</i> , pp.230–231. Thousand Oaks, CA: SAGE Publications, Inc.	Provides a short overview of the assumptions underlying cross-sectional research.
Hall, J. 2008. 'Cross-sectional survey design' in Lavrakas P.J. (ed.) <i>Encyclopedia of Survey Research Methods</i> , pp. 173–174. Thousand Oaks, CA: SAGE Publications, Inc.	Provides a short overview of how to conceptualise, design and conduct a cross-sectional survey.
Liu, C. 2008. 'Cross-sectional data' in Lavrakas, P. J. (ed.) <i>Encyclopedia of Survey Research Methods</i> , pp. 171–173. Thousand Oaks, CA: SAGE Publications, Inc.	Provides a short overview of cross-sectional data in relation to survey studies.
Spector, P. E. 2004. 'Cross-sectional data' in Lewis-Beck, M.S., Bryman, A. and Liao, T.F. (eds.) <i>The SAGE Encyclopedia of Social Science Research Methods</i> , p. 230. Thousand Oaks, CA: SAGE Publications, Inc.	Provides a short introduction to cross-sectional data.
Treiman, D.J. 2009. <i>Quantitative data analysis</i> . San Francisco, CA: John Wiley and Sons, Inc.	Provides an introduction to quantitative data analysis methods, with an emphasis on using the data analysis to draw conclusions in conducting social science research.
Data mining	
Abonyi, J. and Feil, B. 2000. <i>Cluster Analysis for Data Mining and System Identification</i> . Berlin: Birkhauser.	Explains the various steps involved in Knowledge Discovery in Databases, including data mining and the role it plays in the larger process.
Chen, M.S., Han, J. and Yu, P.S. 1996. 'Data Mining: An Overview from a Database Perspective'. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 8(6):866–883.	Provides a survey of the data mining techniques available.
Fayyad, F., Piatetsky-Shapiro, G. and Smyth, P. 1996. 'From Data Mining to Knowledge Discovery in Databases'. <i>American Association for Artificial Intelligence</i> , 17(3):37–54.	One of the core data mining texts – while slightly dated, it provides an informative overview of the subject.

National Research Council. 2008. <i>Protecting Individual Privacy in the Struggle Against Terrorists: A Framework for Program Assessment</i> . Washington DC: National Academies Press.	Talks about the data mining method, amongst others, in terms of countering terrorist threats. While this is slightly beyond an introductory-level text, it offers clear, in-depth information, as well as useful case studies exemplifying the use of data mining.
Saiya, N. and Scime, A. 2014. 'Explaining religious terrorism: A data-mined analysis'. <i>Conflict Management and Peace Science</i> , 32(5):1–26.	Provides basic explanations and more advanced techniques. Focuses particularly on the classification technique.
Descriptive statistics	
Bauer, J. 2009. <i>Statistical Analysis for Decision Makers in Healthcare: Understanding and Evaluating Critical Information in Changing Times</i> , 2nd ed., ch. 5.	Gives an overview of why use each descriptive statistical method, alongside examples within a data set.
Gould, R.N. and Ryan, C.N. <i>Introductory Statistics</i> . (particularly chapters 1, 2 and 3).	Gives an introductory overview, as well as how to interpret findings.
Trochim, W.M.K. 2006. 'Descriptive Statistics.' As of 18 February 2018: http://www.socialresearchmethods.net/kb/statdesc.php	Gives a useful overview of the basic statistical methods used in descriptive statistics and highlights the difference with inferential statistics.
Desk-based research and literature review	
Hague, P. and Wilcock, C. 2014. 'How To Get Information For Next To Nothing'. B2B International. As of 18 February 2018: http://www.b2binternational.com/publications/desk-research/	Focuses on market research, but highlights key features of desk research, and the range of sources available.
Knopf, J. 2006. 'Doing a Literature Review', <i>PS Online</i> , pp. 127–132. As of November 2017: https://www.cambridge.org/core/journals/ps-political-science-and-politics/article/doing-a-literature-review/00B62000B6760AB78E1BD27E32A94C9F	Short paper on how to conduct a basic literature review, which can form part of a desk research exercise.
Focus groups	
Krueger, R.A. 1998. <i>Focus Groups: A Practical Guide for Applied Research</i> . Newbury Park, CA: Sage.	Gives hands-on advice to those seeking to set up and conduct effective focus group interviews.
Liamputtong, P. 2011. <i>Focus Group Methodology: Principle and Practice</i> . London: Sage.	An introductory text that leads readers through the entire process of designing a focus group study.

Morgan, D.L. and Krueger, R.A. 1997. <i>The Focus Group Kit: Volumes 1–6</i> , London: Sage.	An in-depth instruction manual on how to run a successful focus group.
Interviews	
Bryman, A. 2001. <i>Social Research Methods</i> . Oxford: Oxford University Press.	Offers an encyclopaedic introduction to social research methodology, and considers a broad range of qualitative and quantitative methods.
Clutterbuck, L. and Warnes, R. 2013. 'Interviewing government and official sources' in Dolnik, A. (ed.) <i>Conducting Terrorism Field Research: A guide</i> , pp.15–25. Oxford: Routledge.	Provides relevant and useful information for those setting out to interview government, official and practitioner sources on sensitive issues in terrorism and counter-terrorism.
European Commission, Directorate-General for Regional and Urban Policies. 2013. <i>Evalsed Sourcebook: Methods and Techniques</i> . European Commission.	A comprehensive document providing an overview of methods and techniques common in EU evaluations.
National Alliance of Children's Trust and Prevention Funds. 2013. 'Key Informant Interviews (KIIs)'. <i>The Research Savvy Practitioner</i> , 3(1). As of 18 February 2018: http://www.ctfalliance.org/research_savvypractitioner.htm	Provides details on the steps required to conduct key informant interviews.
Roulston, K., deMarrais, K. and Lewis, J.B. 2003. 'Learning to interview in the Social Sciences'. <i>Qualitative Inquiry</i> , 9:643–669.	Provides useful information on how to conduct interviews, and what factors to consider when doing so.
Logic models/theory of change	
Dawson et al. 2014. <i>Learning and Adapting: The Use of Evaluation and Monitoring in Countering Violent Extremism – A Handbook for Practitioners</i> .	Handbook providing some practical guidance on how to elaborate and employ a logic model.
Kellogg Foundation. 2004. 'Logic Model Development Guide'. As of 18 February 2018: https://www.wkkf.org/resource-directory/resource/2006/02/wk-kellogg-foundation-logic-model-development-guide	Provides practical assistance to organisations that aim to demonstrate the effectiveness of their programme activities.
Rogers, P. and Funnell, S. 2011. <i>Purposeful Program Theory: Effective Use of Theories of Change and Logic Models</i> . San Francisco: Jossey Bass.	Handbook setting out the use of logic models and theory of change.

Longitudinal	
Bain, C. and Webb, P. 2005. <i>Essential Epidemiology: An Introduction for Students and Health Professionals</i> . Cambridge: Cambridge University Press.	Any basic epidemiology text book is a good place to start when exploring or thinking about a longitudinal study design for the first time. The following online resources may all also be helpful in particular situations.
Payne, G. and Payne, J. 2004. 'Longitudinal and Cross-sectional Studies' in: <i>Key Concepts in Social Research</i> , pp. 144–149. SAGE Publications, Ltd.	Key text for social research.
Stouthamer-Loeber, M., van Kammen, W. and Loeber, R. 1992. 'The nuts and bolts of implementing large-scale longitudinal studies'. <i>Violence and victims</i> , 7(1):63–78.	Introduction to Longitudinal studies.
Meta-analysis	
Cooper, H., Hedges, L.V., Valentine, J.C. (eds.). 2009. <i>The Handbook of Research Synthesis and Meta-Analysis</i> , 2nd ed., New York, NY: Russell Sage Foundation.	A comprehensive description of the process of undertaking research syntheses.
Lipsey, M.W. and Wilson, D.B. 2001. <i>Practical meta-analysis, Applied social research methods series</i> (vol. 49). Thousand Oaks, CA: Sage.	A step-by-step guide to conducting meta-analysis.
Patton, M.Q. 2014. <i>Evaluation Flash Cards: Embedding Evaluative Thinking in Organizational Culture</i> . St. Paul, MN: Otto Bremer Foundation.	Meta-analysis as a cumulative research technique for programme evaluations.
Pettigrew, T.F. and Tropp, L.R. 2006. 'A meta-analytic test of intergroup contact theory'. <i>Journal of Personality and Social Psychology</i> , 90(5):751–783.	A meta-analytic test of intergroup contact theory.
Network analysis	
Borgatti, S.P. and Molina, J.L. 2003. 'Ethical and Strategic Issues in Organizational Social Network Analysis'. <i>The Journal of Applied Behavioral Science</i> , 39(3):337–349.	Presents the ethical issues that can arise in conducting a social network analysis.
Burnap, P., Williams, M.L., Sloan, L., Rana, O., Housley, W., Edwards, A., Knight, V., Procter, R. and Voss, A. 2014. 'Tweeting the Terror: Modelling the Social Media Reaction to the Woolwich Terrorist Attack'. <i>Social Network Analysis and Mining</i> , 4(1).	Uses Twitter data to build models to predict information flow size and survival after major terrorist events.

Farley, J. 2003. 'Breaking al-Qaeda Cells: A Mathematical Analysis of Counterterrorism Operations (A Guide for Risk Assessment and Decision Making)'. <i>Studies in Conflict and Terrorism</i> , 26(6):399–411.	Uses network analysis to quantify the degree to which a terrorist network is still able to function.
Hanneman, R.A., and Riddle, M. 2005. <i>Introduction to social network methods</i> . Riverside, CA: University of California, Riverside.	Online textbook introducing many of the basics of formal approaches to social network analysis.
Knoke, D. 2015. <i>Emerging Trends in Social Network Analysis of Terrorism and Counterterrorism. Emerging Trends in the Social and Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource</i> , 1–15.	This paper explores the utility and key issues with using network analysis to track transnational terror trends.
Koschade, S. 2006. 'A social network analysis of Jemaah Islamiyah: The applications to counterterrorism and intelligence'. <i>Studies in Conflict and Terrorism</i> , 29(6):559–575.	Presents a social network analysis of communication amongst Jemaah Islamiyah members.
Krebs, V.E. 2002. 'Mapping networks of terrorist cells'. <i>Connections</i> 24(3):43–52.	Explores the difficulties in mapping covert networks and how network analysis can be used.
Ressler, S. 2006. 'Social network analysis as an approach to combat terrorism: past, present, and future research'. <i>Homeland Security Affairs</i> , 2(2):1–10.	Explores the strengths and limitations of network analysis as an approach to P/CVE.
Scott, J. 2012. <i>Social network analysis</i> . London: Sage.	Provides an overview of network analysis, its applications and further details on centrality measures.
Sparrow, M. 1991. 'The Application of Network Analysis to Criminal Intelligence: An Assessment of the Prospects'. <i>Social Networks</i> , 13:251–252.	Explores the opportunities for the application of network analytic techniques to the problems of criminal intelligence analysis, paying particular attention to the identification of vulnerabilities in different types of criminal organisation, from terrorist groups to narcotics supply networks.
Objectives and options analysis	
Fleisher, C. S. and Bensoussan, B. E. 2003. <i>Strategic and Competitive Analysis: Methods and Techniques for Analyzing Business Competition</i> . Upper Saddle River, NJ: Prentice Hall.	Describes the background, rationale and implications, advantages, limitations, process and related tools of objectives analysis.

MethodFinder. 'Problem Tree Analysis'. As of 18 February 2018: http://www.methodfinder.net/briefdescription1.html	An online methods database that describes the main uses, advantages, limitations, and procedures of objectives analysis.
My Management Guide. 2012. 'Project Feasibility and Option Analysis Template'. As of 18 February 2018: http://www.mymanagementguide.com/project-feasibility-and-option-analysis-template/	Outlines how to perform an analysis of feasibility and options that are applicable to a given project design.
Overseas Development Institute. 2009. <i>Problem Tree Analysis</i> . London: Office for International Development.	Provides an outline of objectives analysis.
Observation techniques/ethnography	
European Commission, Directorate-General for Regional and Urban Policies. 2013. <i>Evalsed Sourcebook: Method and Techniques</i> . European Commission. As of 18 February 2018: http://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/evaluation_sourcebook.pdf	Sourcebook outlining a range of methods and techniques for application in the evaluation of socio-economic development.
Kawulich, B.B. 2005. <i>Participant Observation as a Data Collection Method</i> , 6(2): Art. 43. s of 18 February 2018: http://www.qualitative-research.net/index.php/fqs/article/view/466/996#g1	Details the historical roots of participant observation in particular and details its purpose and how a researcher uses this method.
Mills, A.J., Durepos, G. and Wiebe, E. 2010. 'Nonparticipant observation'. <i>Encyclopaedia of Case Study Research</i> . As of 18 February 2018: http://srmo.sagepub.com/view/encyc-of-case-study-research/n229.xml	Sourcebook outlining a range of methods appropriate for case study research.
Policy scientific approach	
European Commission, Directorate-General for Regional and Urban Policies. 2013. <i>Evalsed Sourcebook: Method and Techniques</i> . European Commission. As of 18 February 2018: http://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/evaluation_sourcebook.pdf	Sourcebook outlining a range of methods and techniques for application in the evaluation of socio-economic development, including a practical guide to conducting theory-based evaluations.
Hansen, M.B. and Vedung, E. 2010. 'Theory-based stakeholder evaluation'. <i>American Journal of Evaluation</i> , 31(3)295–313.	Discusses theory-based stakeholder evaluation, a new approach to theory-based evaluation that aims not to create one unitary programme theory but to keep separate the programme theories of diverse stakeholder groups.

Leeuw, F. 2003. 'Reconstructing program theories: methods available and problems to be solved'. <i>American Journal of Evaluation</i> , 24(1)5–20.	Presents three approaches that can be used for reconstructing the programme theories underlying a policy or intervention.
Leeuw, F. 2012. 'Linking theory-based evaluation and contribution analysis: Three problems and a few solutions'. <i>Evaluation</i> , 18(3)348–363.	Provides a list of tools that can be used for argumentation visualisation, and describes one of them, 'rationale', in detail.
Vaessen, J. 2006. 'Programme theory evaluation, multicriteria decision aid and stakeholder values – a methodological framework'. <i>Evaluation</i> , 12(4)397–417.	Focuses on the use of multicriteria decision aid as a tool for eliciting stakeholder values, but also considers some of the drawbacks of using a policy scientific approach.
Qualitative data analysis	
Bazeley, P. and Jackson, K. (eds.). 2013. <i>Qualitative Data Analysis with NVivo</i> . London: SAGE.	A practical introduction to planning and conducting qualitative data analysis with NVivo.
Birks, M. and Mills, J. 2011. <i>Grounded Theory: A Practical Guide</i> . London: SAGE.	Describes the key steps involved in applying grounded theory, from planning a study to presenting its results.
Bryant, A. and Charmaz, K. (eds.). 2007. <i>The SAGE Handbook of Grounded Theory</i> . London: SAGE.	A comprehensive overview of the theory and practice of grounded theory. Covers the method's features, its demands on the skills and capabilities of the researcher, and its position in the domain of research methods.
Flick, U. (ed.). 2014. <i>The SAGE Handbook of Qualitative Data Analysis</i> . Thousand Oaks, CA: SAGE.	The first handbook to provide a state-of-the-art overview of the whole field of qualitative data analysis.
Glaser, B.G. and Strauss, A.L. 1967. <i>The Discovery of Grounded Theory: Strategies for Qualitative Research</i> . New Jersey: Aldine Transaction.	An early work on grounded theory that focuses more on the method's theory than on putting it into practice.
Miles, M.B., Huberman, A.M. and Saldana, J. 2013. <i>Qualitative Data Analysis: A Methods Sourcebook</i> , 3rd edition. London: SAGE.	Describes the fundamentals of qualitative research design and data management.
Quasi-experimental designs	
Shadish, W.R. 2013. Propensity score analysis: promise, reality and irrational exuberance. <i>Journal of Experimental Criminology</i> , 9:129–144.	Provides an overview of propensity score matching and analysis, including its potential and limitations.

Shadish, W.R. and Clark, M.H. 2004. 'Quasi-experiment' in Lewis-Beck, M.S., Bryman, A. and Liao, T.F. (eds.). 2004. <i>The SAGE Encyclopedia of Social Science Research Methods</i> , pp. 899–903. Thousand Oaks, CA: SAGE Publications, Inc.	Provides an overview of quasi-experimental designs, including their historical development, historical examples of quasi-experimental designs, and the statistics used to improve causal inferences.
Shadish, W.R., Cook, T.D. and Campbell, D.T. 2002. <i>Experimental and quasi-experimental designs for generalized causal inference</i> . Boston: Houghton-Mifflin.	Provides information on conducting experimental and quasi-experimental research with a focus on the elements, challenges and solutions of research design.
Trochim, W.M.K. 2006. 'Quasi-experimental design'. <i>Research Methods Knowledge Base</i> . As of 18 February 2018: http://www.socialresearchmethods.net/kb/quasiexp.php	Provides an introduction to quasi-experimental designs, including specific descriptions of the non-equivalent groups design and the regression-discontinuity design.
White, H. and Sabarwal, S. 2014. 'Quasi-experimental design and methods'. UNICEF Methodological briefs, Impact Evaluation No. 8. As of 18 February 2018: http://devinfo.liveinfo.org/impact_evaluation/ie/img/downloads/Quasi-Experimental_Design_and_Methods_ENG.pdf	Provides insight into quasi-experimental design and methods for evaluation purposes.
Randomised control trials	
CONSORT n.d. 'Welcome to the CONSORT Website'. As of 18 February 2018: http://www.consort-statement.org/	A website that sets out best-practice reporting standards for RCTs and provides checklists.
Haynes, L., Service, O., Goldacre, B., and Torgerson, D. 2012. 'Test, Learn, Adapt: Developing public policy with randomised control trials'. Cabinet Office: Behavioural Insights Team. As of 18 February 2018: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/62529/TLA-1906126.pdf	An introduction to RCTs aimed at policymakers and others in government interested in evidence-based policymaking.
HM Treasury. 2011. <i>The Magenta Book: Guidance for evaluation</i> . As of 18 February 2018: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/220542/magenta_book_combined.pdf	Guidebook offering in-depth guidance on how evaluations should be designed and undertaken.
Shadish, W.R., Cook, T.D. and Campbell, D.T. 2002. <i>Experimental and quasi-experimental designs for generalized causal inference</i> . Boston: Houghton-Mifflin.	Provides information on conducting experimental and quasi-experimental research with a focus on the elements, challenges and solutions of research design.

Teoh, P.J. and Camm, C. F. 2013. 'Test, learn, adapt'. <i>Annals of Medicine and Surgery</i> , 2(1):22–25. doi: 10.1016/S2049-0801(13)70023-2	An open-access journal article summarising the Test, Learn, Adapt publication and discussing some elements in more detail.
Realist evaluation	
Better Evaluation n.d. 'Realist evaluation'. As of 18 February 2018: http://betterevaluation.org/approach/realist_evaluation	Provides an introduction to realist evaluation and includes links to other resources, including an introduction to realist impact evaluation.
Gielen, A.J., 2017. Countering Violent Extremism: A Realist Review for Assessing What Works, for Whom, in What Circumstances, and How?. <i>Terrorism and Political Violence</i> , pp. 1-19. As of 18 February 2018: https://www.tandfonline.com/doi/abs/10.1080/09546553.2017.1313736	Argues that we should move away from the "what works?" question and towards: "what works, for whom, in what circumstances, and how?".
Pawson, R. and Tilley, N. 2004. <i>Realist evaluation</i> . As of 18 February 2018: http://www.communitymatters.com.au/RE_chapter.pdf	This chapter provides an overview of conducting realist evaluations, including an explanation of basic concepts, methods, and strengths and limitations of the approach.
University of Liverpool, Centre for Advancement in Realist Evaluation and Synthesis (CARES). n.d. 'Useful Resources'. As of 18 February 2018: https://www.liv.ac.uk/psychology-health-and-society/research/cares/realist-videos/	Provides a series of videos of an interview with Ray Pawson (who together with Nick Tilley developed the realist evaluation approach) and a link to other introductory resources on realist evaluation.
Stakeholder analysis	
Better Evaluation. 2014. 'Stakeholder mapping and analysis'. Better Evaluation website. As of 18 February 2018: http://betterevaluation.org/evaluation-options/mapping_stakeholders	A brief introduction to stakeholder mapping and analysis with a step-by-step guide.
Borgatti, S.P. and Molina, J.L. 2003. 'Ethical and Strategic Issues in Organizational Social Network Analysis'. <i>The Journal of Applied Behavioral Science</i> , 39(3):337–349.	Presents the ethical issues that can arise in conducting a social network analysis.
Nyheim, D., Leonhardt, M. and Gaigals, C. 2001. 'Development in conflict: A seven step tool for planners'. London: FEWER-International Alert-Saferworld. As of 18 February 2018: http://www.dis.uniroma1.it/~webemer/Dispense/Pietrantonio/devinconflict.pdf	Draws on a range of sources and expertise and it is based on lessons learned from the conflict research.

World Bank. Stakeholder Analysis. n.d. As of 18 February 2018: http://www1.worldbank.org/publicsector/anticorrupt/PoliticalEconomy/stakeholderanalysis.htm	Excellent resources on stakeholder analysis.
Surveys	
Dillman, D.A. 1978. <i>Mail and Telephone Surveys: The Total Design Method</i> . New York: John Wiley and Sons.	Outlines the comparative advantages and disadvantages of mail, telephone and face-to-face surveys, and describes how to implement each survey type.
Fink, A. 2013. <i>How to Conduct Surveys: A Step-by-Step Guide</i> . London: SAGE publications.	An instruction manual to help researchers to develop rigorous surveys.
Rea, L.M. and Parker, R.A. 2005. <i>Designing and Conducting Survey Research: A Comprehensive Guide</i> , 3rd ed. San Francisco, CA: John Wiley and Sons.	Describes the steps required to design and deploy surveys.
Shonlau, M., Fricker, R.D. and Elliott, M.N. 2002. <i>Conducting Research Surveys via E-mail and the Web</i> . Santa Monica, CA: RAND Corporation, MR-1480-RC.	Provides a useful overview of online surveys.