A literature review on methodology used in evaluating effects of preventive and de-radicalisation interventions

Feddes, A.R.; Gallucci, M.

Published in:
Journal for Deradicalization

Citation for published version (APA):
A Literature Review on Methodology used in Evaluating Effects of Preventive and De-radicalisation Interventions

By: Allard R. Feddes¹
    Marcello Gallucci²

Authors note

This research was made possible by an FP7 European Union research grant (nr. 312235) and is part of a larger research project named IMPACT (Innovative Method and Procedure to Assess Counter-Violent Radicalisation Techniques in Europe, see also: http://impacteurope.eu) including 12 partner organisations across Europe and beyond. Correspondence regarding this article should be addressed to Allard R. Feddes (a.r.feddes@uva.nl), University of Amsterdam, Department of Psychology, Weesperplein 4, 1018 XA, Amsterdam, the Netherlands, or Marcello Gallucci (marcello.gallucci@unimib.it), University of Milano-Bicocca, Department of Psychology, Piazza Ateneo Nuovo 1, Ed. U6, 20126, Milan, Italy.

¹ University of Amsterdam, the Netherlands
² University of Milano-Bicocca, Italy
Abstract
Policies aimed at preventing (further) radicalisation or aimed at de-radicalisation are required to be ‘evidence-based’. This suggests that evaluators should apply rigorous empirical methodology and measurement techniques. However, it is often unclear what this evidence should consist of and how it should be gathered. In the present paper we present results of a literature review focusing on evaluations of programmes aimed at preventing radicalisation or de-radicalisation between 1990 until July 2014. We identified 55 manuscripts including 135 participant samples. Primary qualitative or quantitative empirical data about effectiveness of an intervention was presented in only 16 participant samples (12%). The outcomes are discussed with respect to methods and interventions used in the research field of criminology, a valuable source of methodological experience in conducting evaluation research in challenging circumstances. We recommend the use of empirical studies using quantitative data when possible (i.e., in preventive interventions) and a multi-method approach for evaluating programmes in (even) more challenging contexts (i.e., de-radicalisation programmes).
A Literature Review on Methodology used in Evaluating Effects of Preventive and De-radicalisation Interventions

Recently there has been a strong increase in interest among policy makers, first-line workers, and researchers in the question "what works" to prevent radicalisation or to de-radicalise individuals so that they are unwilling to use violence to reach their ideals. For example, in Europe, there has been a call for more attention to prevent radicalisation (Communication from the European Commission, 2014) and the EU Counter-Terrorism Coordinator explicitly called for a systematic investigation in "lessons learned, good practices, unsuccessful practices, and analyse why certain approaches have succeeded or not, in order to develop expertise on what makes for successful interventions" (Council of the European Union, 2011, p. 6). Indeed, this has resulted in initiatives to collect and compare interventions aimed at countering radicalisation such as the Radicalisation Awareness Network (RAN; 2014).

Radicalisation can be understood as a process of increasing willingness to use violence in order to obtain political or religious goals (see also Slootman & Tillie, 2006; Horgan & Braddock, 2010). In line with this description, Horgan defines de-radicalisation as "the social and psychological process whereby an individual’s commitment to, and involvement in, violent radicalisation is reduced to the extent that they are no longer at risk of involvement and engagement in violent activity" (2009, p. 153).

Even though a range of programs aimed at prevention of radicalisation and de-radicalisation have been designed and implemented, the impact of these programs, the underlying mechanisms involved, as well as economical costs are often not clear. Previous systematic reviews in regard to counter-radicalisation interventions have been limited to
identifying the effect of government strategies (Lum, Kennedy, & Sherley, 2006) and examinations of process(es) of (de-) radicalisation and available interventions to prevent radicalisation (Christmann, 2012; Demant, Slootman, Buijs, & Tillie, 2008). Lum and colleagues concluded that there was an “almost complete absence of evaluation research on counter-terrorism interventions” (p. 489). The more recent systematic review by Christmann did not show signs of improvement as he concluded that the “evidence base for effective preventing violent extremism interventions is very limited” and “despite a prolific output of research, few studies contained empirical data or systematic data analysis”.

Nevertheless, counter-radicalisation policies in countries like Denmark, Germany, the Netherlands, the USA and the UK are required to be ‘evidence-based’ which suggests that evaluators should apply rigorous empirical methodology and measurement techniques. However, it is often unclear what this evidence should consist of and how it should be gathered (see also Bovenkerk, Van Hemert, & Quint, 2013; Feddes, Mann, & Doosje, 2013, 2015; Gielen, 2015; Horgan, 2009; Köhler, 2013). Also, there is to date no consensus on indicators of successful de-radicalisation (Barrett & Bokhari, 2009; Horgan & Braddock, 2010; Vidino, 2010). It has been posed, therefore, that interventions that have been evaluated often do not meet scientific standards (Bovenkerk et al., 2013; Carline, 2011; Christmann, 2012; Dalgaard-Nielsen, 2010; Horgan & Braddock, 2010; Lindekiilde, 2012; Lub, 2013; Lum et al., 2006).

Even though these research gaps have been identified, to date no systematic overview exists that outlines what methods and techniques have actually been used to assess interventions and evaluate impact of counter-radicalisation interventions. This information is important when considering future policy to evaluate these interventions. The aim of the present paper is to outline the methods used thus far.
Method

Data sources

A systematic literature search was conducted in the period May until July 2014 using three strategies: (1) we sent a direct request for (un)published manuscripts to 45 researchers and experts in the field from countries including Australia, Denmark, France, Germany, Italy, the Netherlands, Norway, Spain, the UK, and the USA; (2) an online literature search was performed using a series of keywords in a selection of online databases including PsycINFO, PUBMED, COCHRANE Library, WEB of SCIENCE, ERIC, SCIENCE DIRECT, the National Criminal Justice Reference Service (NCJRS, USA), and the UK Home Office Research Database; (3) we used the so-called “snowball method” to find additional possible relevant manuscripts by examining the reference lists of manuscripts that were considered suitable to include in the review.

Search strategy

The search terms and the results of the online literature search are given in Table 1. As can be seen, the number of hits ranged from zero to 2,591 (a background search was conducted first using Google Scholar resulting in unmanageable numbers of hits; 17,500 or more). The primary reviewer first checked the titles and abstracts for relevance. In case a manuscript was deemed relevant but could not be accessed, the authors were approached directly.
Inclusion and exclusion criteria

A manuscript was included when a study was conducted in which an intervention was evaluated aimed at preventing (further) radicalisation or aimed at de-radicalisation. A second criterion was that the intervention was evaluated using a qualitative or quantitative evaluation method. It was decided to only include manuscripts reporting evaluations of interventions from 1990 onwards until July 2014.

Coding Procedure

A coding scheme and detailed coding instructions (available from the authors on request) were constructed. Both authors (who have a behavioural science background) coded the manuscripts. Three manuscripts were coded together and possible disagreements were discussed until agreement was met or alterations in the coding scheme were made. The remaining manuscripts were coded by the first author and in case of doubt the coders again discussed until agreement was met.

Results

Based on the two criteria mentioned above, ultimately 55 manuscripts met the three criteria given above. These manuscripts are identified in the reference list with an "*". Nine of the 55 manuscripts (16%) came from the request that had been sent to the authors by colleagues in the field. The remaining 46 samples were found by means of the online search or the snowball method.
As can be seen in Figure 1, the majority of manuscripts (88%) came from 2008 or later. Forty-seven of the manuscripts (85%) were published, however, only 13 of these (28% of the total number of manuscripts included in the review) had been peer reviewed. As can be seen in Table 2, the majority of manuscripts reported on studies conducted in, respectively, the Netherlands, the UK, Saudi Arabia, Denmark, and Germany. A relatively large number came from the Netherlands. This is partly due to the fact that we had a relatively large social network of researchers and practitioners in the Netherlands.

Description of sample and interventions

Data were structured on basis of ‘Intervention Focus’. Different foci were distinguished; an intervention could have been an individual or group who/which, in terms of radicalisation is non-radical (having shown no interest in an extremist ideology), potentially radicalising (having shown an interest in but not having shown extremist behaviour) or radicalised (behaviour related to an extremist ideology has been observed). Typical examples of interventions focusing on non-radical groups are educational interventions or workshops that aim at adolescents and young adults. For example, the effect of an interactive exhibition to promote knowledge about democracy in the Netherlands (the so-called “Fortress of Democracy”) was investigated by Van Ooijen (2011) and Huijzer (2012).

An example of an intervention focusing on both potentially radicalising and radicalised individuals is the De-radicalisation Targeted Intervention project by the Danish Ministry of Social Affairs and Integration (Korf, 2012; COWI, 2014). The Danish strategy is characterized by a focus on, respectively, young persons who are considered vulnerable to
extremism and extremist individuals. Individuals considered to be potentially radicalising are described as showing a fascination for or sympathy with extremist environments, form of actions, and/or ideologies, and those who are considered to be shortly involved in extremist environments. Radicalised individuals are described as being integrated in or actively participating in extremist environments or having been convicted for the "terror act".

An intervention could also have focused on the social context surrounding these primary targets groups. This social context includes the community, family and friends, and first-line professionals who work with the target groups (i.e., social workers, police). A good example of an intervention focusing on the community is the violent extremism prevention programme by the Muslim Council of Wales. This programme includes awareness raising of radicalisation in the community, training of community members and English lessons for Imams (see Braga & Weisburd, 2012). A focus group that was part of the CRIME study by Lousberg, Griffioen-Young, Dyevre, and Goetz (2010) pointed out that friends and family could serve as protective factors according to experts. Finally, an example of interventions focusing on first-line professionals is the Philippine government’s de-radicalisation efforts where prison personnel are educated in recognizing radicalisation by creating awareness about radicalisation threat (Jones & Morales, 2012). Based on this distinction, 135 different population samples were derived from the 55 manuscripts in the review.

**Intervention focus and ideology.**

In Table 3 the number of samples is given based on intervention focus and the ideology as reported in the respective manuscript. Radicalised violent individuals were mostly focused on (n = 50 samples; 37%), followed by potentially radicalising individuals (n
non-radical groups \((n = 17; 13\%)\), non-radical individuals \((n = 13; 10\%)\), and radicalised violent groups \((n = 10; 7\%)\). The remaining target samples consisted out of seven community samples and one sample of family and friends. Taken together, these results show that when considering non-radical, potentially radical, and radicalised groups versus individuals, most attention in the evaluation studies included in the review was aimed at the individual level of research instead of a group level.

**Intervention goal.**

The results on the goal of evaluated interventions are given in Table 4. In coding intervention goals a distinction was made depending on phase of radicalisation (i.e., preventative, suppressive, or restorative) and time (i.e., short-term and long term). Short-term preventative was coded when the intervention focused on non-radicalized individuals or groups with the aim to book results on a short notice (a period of maximum one month). Two samples were considered short-term preventative. The earlier mentioned violent prevention programme by the Muslim Council of Wales is an example (Braga & Weisburd, 2012). Long-term preventative was coded most often \((n = 62; 46\%)\). This category refers to non-radicalized individuals or groups and interventions that aim to prevent radicalisation on the long term. Examples are programs such as in the UK (Secretary of State for the Home Department, 2011) and the earlier mentioned Danish Prevent approach. Both focus on individuals and groups considered vulnerable for radicalisation (COWI, 2014).

Short-term restorative was coded when individuals or groups had shown an interest in, or already joined an extremist group. In addition, the intervention aimed to make individuals leave the group (disengage) or de-radicalise on short notice (within a period of
weeks until 1 month). In four samples a short-term restorative intervention was encountered. An example is the Danish Prevent approach aimed at individuals who already showed signs of radicalisation (COWI, 2014).

*Long-term restorative* was coded when the intervention had the goal of making individuals leave the group or de-radicalise. In addition, the programme should cover a period longer than a month. This was the case in 49 samples (36%). An example is the Saudi de-radicalisation programme that aims to de-radicalise convicted extremists in prison and provides long-term support to participants in the programme as well as their relatives (see Boucek, 2007, 2008a, 2008b). *Short-term suppressive* was coded when the intervention aimed to suppress ideology-based violence and behaviour over a short period of time (less than 1 month) by means of, for example, policing work. No samples were encountered using this method.

*Long-term suppressive* was coded when intervention aimed to suppress ideology-based violence and behaviour over a longer period of time (more than one month). Eleven per cent of the samples was found to concern long term suppression. An example is the Jordan approach towards violent extremism that mainly relies on suppressive measures like infiltration in suspected groups, arrests and imprisonment (El-Said, 2012). All in all, it can be concluded from these data that counter radicalisation interventions mainly had long-term goals. As will be seen below, this is problematic as most of the research methods used do not take a long-term approach into account.
Findings related to evaluations

Evaluation factors provide information about the assessment method and instruments used in the study. Below, detailed information is given how these factors were coded and the findings are presented.

Quality of data.

As reported above, whereas the majority of counter-radicalisation samples came from published manuscripts, most of these had not been subjected to peer review. As an indicator of source quality it was coded whether the data was anecdotal, empirical, or theoretical. Anecdotal was coded when the manuscript provided a description of the intervention but this was not related to any theory and no empirical data was collected. It was found that the majority of samples belonged to this category ($n = 66; 49\%$). Theoretical was coded when a theory was tested by means of a review of the literature but no qualitative or quantitative data was collected to test the hypotheses. A total of 53 samples (39\%) came from manuscripts that had a theoretical focus. Finally, empirical was coded when quantitative or qualitative data was collected to examine the impact of the intervention. It was found that only 16 samples (12\%) came from interventions that had been empirically evaluated reporting primary data. An example of such an evaluation is an evaluation conducted by Lousberg, Van Hemert, and Langelaan (2009) who used an online questionnaire for first-line workers to evaluate effectiveness of interventions. Another example is a study of effectiveness of de-radicalisation programme of Liberation Tigers of Tamil Eelam (LTTE also referred to as the "Tamil Tigers") in Sri-Lanka by Kruglanski,
Bélanger, Gelfand, Gunaratna, and Hettiarachchni, (2014) including 1,906 participants. A combination of interviews and surveys was used to investigate effectiveness of this programme.

An example of an evaluation that was not evaluated by means of collecting empirical data is the Saudi rehabilitation programme described by Boucek (2008a, 2008b). Boucek emphasises that understanding the Saudi rehabilitation programme aimed at Islamist extremists and militants is important as it is a program that is "the best funded and longest continually run programme" (2008a, p. 64-65). According to this author the programme, which is characterised by extensive social support given to detainees and families, is possibly a model for counter-radicalisation programmes for Western partners such as the UK. All in all, these findings illustrate that primary empirical data is relatively scarce and evaluations of interventions are mainly of an anecdotal nature.

Evaluation focus.

In regard to evaluation focus a distinction was made between Impact, Mechanism, Process and Economic focus. Impact was coded when the evaluation examined the effect of the intervention. Mechanism was coded when the underlying mechanism was of interest, that is, the evaluation focused on the question why the intervention was considered effective (or not). Process was coded when an assessment was made in regard to how the programme was implemented. That is, whether or not elements of the programme were successfully implemented (or not implemented at all). The option Economic was coded when financial costs were considered. As can be seen in Table 5, most evaluations concerned a combination of Impact and Mechanism (n = 62; 46%), followed by a combination of Impact, Mechanism
as well as Process evaluation \((n = 31; 23\%)\). Economic aspects of the intervention like the costs of the intervention were included in samples \((19\%)\). These include a study by Demant, Slootman, Buijs and Tillie \((2008)\) who took into account the approximate costs when considering restorative interventions with right-wing extremists. Another example is the review of effects of the Prevent programme in which costs were considered as well \((\text{Department for Communities and Local Government, 2008})\).

In a next step we examined focus of interventions across intervention population sample \((\text{i.e., individual non-radical, group radicalised})\). It was found that evaluations of programmes aimed at potentially radical individuals and radicalised individuals, most often focused on a combination of impact and underlying mechanism. In the majority of samples \((n = 119; 88\%)\) the underlying mechanism \((\text{why does an intervention work or not})\) was considered.

**Evaluation method and instruments.**

From a social science perspective, a rigorous scientific measurement of the impact or outcome of an intervention involves the use of methods and instruments to collect empirical data. This data should allow for an objective comparison of ‘what works’. We make a distinction between methods and instruments.

A range of methods were coded in the present review: *Experimental* was coded when the intervention included an experimental and control group and the researcher controlled assignment of participants to the experimental and control group; a *Quasi-experimental* method was coded when there was an experimental group but there was no control over who was assigned to experimental/control group or the control group was not
present; a Longitudinal with follow-up was coded when data was collected at multiple points in time, but at least a pre- and post-measurement needed to be present with a follow-up measurement later in time (at least 1 month later). Longitudinal without follow-up was coded when no follow-up measurement took place. Cross-sectional was coded when data was collected only at one point in time (e.g. only a post-measurement). Cross-historical comparison was coded when interventions at different points in time are compared (e.g., two interventions are compared which had been conducted in the 1970s and 1990s). A Case study was coded when the evaluation focused on a specific individual or group, or a specific event. Meta-analysis was coded when results from different studies were combined and analysed using statistical methods.

In Table 6 an overview is given of the number of times (combinations of) methods have been used in counter-radicalisation interventions. It was found that for 74 samples (55%) a cross-sectional evaluation was applied. For example, a cross-sectional evaluation was conducted by the above-mentioned Lousberg and colleagues (2009). A small minority of assessment studies applied a longitudinal design (n = 5; 4%) of which only one study focusing on a non-radical sample included a follow-up measurement. In this study a group of non-radical individuals were interviewed before, during, and after participating in a training program which was aimed at increasing resilience (Feddes et al., 2013). Quasi-experimental methods were used in only three samples. A noteworthy evaluation study using a quasi-experimental design including experimental and control groups, is the before-mentioned evaluation of the reintegration program of Tamil Tigers in Sri-Lanka by Kruglanski et al. (2014). Noteworthy is that in 50 of the included 135 samples (37%), the evaluation method was not specified. As can be seen in Figure 2, the use of multiple methods to evaluate effectiveness was found in not more than four samples (3%).
In Table 7 an overview is given of instruments which have been used divided across intervention sample. Instrument here refers to a technique (e.g., a Focus group), a device (e.g., a Questionnaire) or a research process (e.g., Data mining, meta-analysis) used to gather information regarding the intervention to be evaluated. For example, Quantitative survey was coded when respondents in the study completed a questionnaire resulting in a quantitative dataset. In the majority of evaluations no empirical instruments were specified ($n = 54$, 40%, see also Figure 3). Observation was most often used to investigate effectiveness in itself or in combination with other instruments ($n = 41$, 30%). Qualitative interviews were used second often ($n = 37$, 27%). Observations were used, for example, in evaluations of restorative interventions with radicalised individuals participating in the EXIT programs (i.e., BJØRGO & CARLSSON, 2005; THOMSEN, 2012) but also in evaluation of preventative studies (i.e., KPLUSV, 2010). In 28 samples (21%) multiple instruments were used. An example is the evaluation of a training for first-line workers to counter radicalisation in Amsterdam (PELS, 2009). Three trainings given by a research company were evaluated by means of observations by two independent researchers and the use of questionnaires that were sent to participants in the course.

Theory-based approach.

Practical restraints often make it difficult to collect sufficient empirically based data for analyses of effects and comparisons of evaluations. CHEN and ROSSI (1983) argued for a theory-driven approach to compensate for shortcomings of research designs that do not meet the high standards of a randomized controlled experimental design. Indeed, it is considered as good practice in the evaluation field to have a so-called ‘Theory of Change’ that makes
explicit the different components of the intervention and the outcome of each component, as well as the expected relations between the different components. This descriptive element of an intervention should result in testable hypotheses that can then be answered by means of collecting data using qualitative or quantitative measurement techniques or a combination of both (Leeuw & Martini, 2013; Lindekiilde, 2012; Lub, 2013).

We coded different theory-based evaluation approaches. In most samples no theory-based evaluation was specified ($n = 81; 60\%$). As can be seen in Table 8, a so-called *policy scientific approach* was coded in 34 samples (25%). This approach involves identifying the behavioral systems (mechanism) expected to counter radicalisation and link these with policy programmes (Leeuw & Martini, 2013). An example is the previously mentioned evaluation conducted by COWI (2014).

*A theory of change* refers to the procedure of describing assumptions that explain the steps leading to a long term goal as well as connections between programme activities and outcomes at each step of the way (Weiss, 1995). We coded a theory of change in 16 samples (12%). One example is the evaluation of preventative interventions in the Netherlands (KplusV, 2010; Pels, 2009). *A contribution analysis* (a measure of performance that aims to establish the contribution a programme makes to desired outcomes; Mayne, 2008) was coded in one sample (Gemeente Weert, 2010) and can also be found in an evaluation study focusing on crime by Krafchik (2011). In another sample (Sheikh, Sarward, & King, 2012) a *realist evaluation approach* was coded. This approach stresses the importance of context, mechanisms involved, and outcomes to learn more about ‘what works for whom’, ‘in which context does a programme work’, ‘what mechanisms are triggered’ (see Pawson & Tilley, 1997; see Lobley, Smith, & Stern, 2001, for an example in context of criminal behaviour).
Conclusions and Recommendations

In addition to previous reviews, the present review shows that up to July 2014 hardly any empirically based evidence of preventive or de-radicalisation interventions exist. By means of a systematic coding procedure we have described 135 samples in terms of intervention goal, evaluation focus, evaluation method, evaluation instruments, and theory-driven approaches that have been used.

Perhaps not surprisingly, the manuscripts included in the review were mostly anecdotal in which no explicit reference to theory and no empirical quantitative or qualitative data was reported. Primary quantitative or qualitative data was presented in only 16 out of 135 samples (12%). Instruments and methods used were often not specified. Cross-sectional methods have been used most often. This is problematic as many interventions have a long-term prevention or restoration focus. In addition, it was found that when used, evaluations often make use of a single instrument. An additional finding is that existing interventions mainly focus on the individual level whereby attention for effects on a group level is lacking.

It is acknowledged that evaluation research of counter-radicalisation interventions is subject to great challenges ranging from pragmatic issues to ethical considerations. Noteworthy, related research fields like criminology could offer insights on how to conduct empirically-based evaluation research nevertheless. This field has over 50 years of experience in evaluating effects of interventions preventing gang involvement and related criminal behaviour in the USA and EU. In this field multi-method quantitative approaches are frequently used (see for a recent review and a recent meta-analysis: Gravel, Bouchard, Descormiers, Wong, & Morselli, 2013; Koehler, Lösel, Akoensi, & Humphreys, 2013).
Indeed, a comparison between the field of radicalisation studies and criminology studies has been made before (e.g., Lindekiilde, 2012; Mullins, 2010). Mullins has investigated rehabilitation programs for Islamist militants in light of rehabilitation interventions for "ordinary" criminal offenders. He acknowledges that many of the obstacles in regard to evaluation of effectiveness (lack of clarity about concepts, difficulties in coming up with viable methodologies for research and evaluation, lack of data) are amplified with research on terrorism.

High-quality evaluations can be encountered in the field of criminology focusing on interventions countering criminal gangs. These studies are characterized research methods including experimental and quasi-experimental counterfactual designs, longitudinal designs, and theory-based evaluations. The evaluations include hypotheses and assumptions that can be empirically tested and often include multiple instruments (interviews, surveys, observations, calculation of indicators of recidivism).

Besides using a wide range of methods and a theory-based approach is considered good practice in evaluating interventions (see Leeuw & Martini, 2013). These would allow for making explicit the underlying assumptions of an intervention and provide goals that could be tested. This approach has been recommended by several researchers. For example, Horgan and Braddock (2010) propose a systematic approach for planning and evaluating de-radicalisation programs. Their view is much in line with Lub (2013) who describes a theory of change approach to evaluate interventions aimed at prevention or de-radicalisation. The present review shows that 54 of the 135 (40%) included samples used this approach as part of their evaluations. This is less than half of all samples which illustrates that this approach is not yet widely used.
Williams and Kleinman (2014) and Leeuw and Martini (2013) stress that theories of change should be used alongside empirical data (quantitative or qualitative) to measure the degree program goals are met and programs are implemented as planned. In contexts that allow for data collection among participants, such as interventions focusing on effectiveness of countering radicalisation among non-radical individuals (i.e., preventive interventions in schools), an empirical approach in which an experimental group is compared to a control group is recommended. However, the evaluation of effectiveness of the reintegration program of Tamil Tigers in Sri Lanka (Kruglanski et al., 2014) in which experimental and control were used, illustrates that empirical data can be collected directly from participants where sufficient. Future research could disentangle the necessary conditions facilitating the use of these methods.

If primary empirical data cannot be collected directly, for example in case of radicalised individuals, an indirect approach could be taken by including the social context (peers, family members, first line workers) in the evaluation. The data methods used to examine change over time could also be collected by measuring objective goals (i.e., finding work, getting back to school). Besides these objective measures a range of methods are available such as mere observation, focus groups, interviews (with the target if possible, otherwise with the social context) or a combination of these. Recent publications by Möller, Küpper, Buchheit, and Neuscheler (2015) and Schuurman and Bakker (2015) nicely illustrate this approach.¹

Möller et al. (2015) report an evaluation study of an EXIT programme in Germany. This programme deals with 'difficult' cases characterized by individuals with criminal records and behavioural misconduct. In this evaluation study, the researchers used a mixed-method approach. The impact of the programme was examined through quantitative data
made available by the organisation running the programme. These data included basic information such as the number of completed cases, the number of current cases, and number of cases that were ended before the programme was finished.

In addition, a process evaluation was conducted to examine in detail the hypotheses underlying the program. The process evaluation involved (1) an analysis of the documents describing the program, (2) an analysis of the impressions of the first-line workers guiding the individuals (six first-line workers were interviewed for this purpose using semi-structured interviews), (3) interviews were conducted with the head of the programme and the spokesperson (also using semi-structured interviews), (4) interviews with former clients about the support they received during the program (eight semi-structured interviews), (5) interviews with people from the direct social context of the client (i.e., two sets of parents in two separate interviews). Interview data were analysed using content analyses focusing on the life situation at different stages, individuals' motivation, and possible factors influencing the de-radicalisation process. The process evaluation, therefore, focused on how the programme was implemented by directly comparing the process to existing documentation of the programme. By focusing on both the documents describing the programme, the underlying processes involved, and connecting this to the expected outcomes, a theory of change was made explicit making an evaluation possible despite the challenging conditions.

Schuurman and Bakker (2015) also took a theory of change approach by first clarifying the underlying assumptions of the Dutch initiative. Following, the author’s conducted a process and impact evaluation by examining whether the programme was implemented as planned and whether the underlying assumptions were correct. Three rounds of semi-structured interviews were held for this purpose. The impact evaluation, in turn, focused on the programme goals: (1) whether recidivism among extremist and terrorist
offenders was reduced; (2) by monitoring (former) extremists and terrorists through mandatory probation and (3) by evaluating a prevention-focused addition to the Dutch authorities’ counterterrorism toolbox. By taking a multi-method approach these researchers were able to disentangle the theory of change and connecting underlying program assumptions to outcomes. This process evaluation made possible a subsequent impact evaluation. Taken together, this allowed for an interpretation of the effect of the programmes despite the challenging conditions that are typical for de-radicalisation programmes.

We see these examples as promising approaches in evaluating de-radicalisation programmes while strongly encouraging conduct of strong empirical studies in circumstances that allow for it. A theory based approach complemented with appropriate methods and instruments for measuring impact will help developing a stronger basis for future policy and programmes aimed at prevention and de-radicalisation.
Footnote

1 We thank two anonymous reviewers who pointed out both studies. The studies were published after the literature review was conducted and it was decided not to include them in the dataset but discuss them instead.
References


*Bjørgo, T. & Carlsson, Y (2005). Early intervention with violent and racist youth groups. NUPI.*


International Peace.


Communication from the European Commission, (2014). Preventing radicalisation to terrorism and violent extremism: Strengthening the EU’s response. *Communication from the commission to the European Parliament, the Council, the European*
Economic and Social Committee and the Committee of the Regions. Retrieved on June 1, 2014 from: http://ec.europa.eu


*De Graaf, B. (2009). Counter narratives and the unrehearsed stories counter-terrorists unwittingly produce. Perspectives on Terrorism, 3, 5-11


Studies in Conflict & Terrorism, 35:3, 211-228.


Research Report by Inspira Consulting.


Amsterdam Muslims become radicals. University of Amsterdam: Institute for Migration and Ethnic Studies.


*Wagner, R. (2010). Rehabilitation and deradicalisation: Saudi Arabia’s
counter-terrorism successes and failures. *Peace and Conflict Monitor*.


Figures

Figure 1.

*Number of Published and Unpublished Manuscripts included in the Review from 1997 onwards*

(N = 55)
Figure 2.

*Number of Evaluation Samples (N = 135) in which the Method used is not Specified, One Method is used, or Multiple Methods have been used*
Figure 3.

Number of Evaluation Samples (N = 135) in which the Instrument used is not Specified, One Instrument is used, or Multiple Instruments have been used

![Bar chart showing the number of evaluation samples for unspecified, one instrument, and multiple instruments](chart_image)
Tables

*Table 1.* Overview of the results of the online search for manuscripts reporting on an evaluation of an intervention focusing on counter-radicalisation published between 01/01/1990 and 01/07/2014

<table>
<thead>
<tr>
<th>Database</th>
<th>PsycINFO</th>
<th>PUBMED</th>
<th>COCHRANE library</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keywords</td>
<td>radicalis</td>
<td>terroris</td>
<td>radicalis</td>
</tr>
<tr>
<td></td>
<td>ation</td>
<td>m</td>
<td>ation</td>
</tr>
<tr>
<td></td>
<td>AND</td>
<td>AND</td>
<td>AND</td>
</tr>
<tr>
<td></td>
<td>interv</td>
<td>interv</td>
<td>interv</td>
</tr>
<tr>
<td></td>
<td>tion</td>
<td>tion</td>
<td>tion</td>
</tr>
<tr>
<td></td>
<td>AND</td>
<td>AND</td>
<td>AND</td>
</tr>
<tr>
<td></td>
<td>evaluati</td>
<td>evaluati</td>
<td>assessme</td>
</tr>
</tbody>
</table>

PsycINFO

http://psycnet.apa.org/

PUBMED


COCHRANE library

Allard Feddes & Marcello Gallucci: A Literature Review on Methodology used in Evaluating Effects of Preventive and De-radicalisation Interventions
<table>
<thead>
<tr>
<th>Database</th>
<th>Pages</th>
<th>Search Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>WEB OF SCIENCE</td>
<td>3</td>
<td>45</td>
</tr>
<tr>
<td>SCIENCE DIRECT</td>
<td>120</td>
<td>1.877</td>
</tr>
<tr>
<td>National Criminal Justice Reference</td>
<td>26</td>
<td>5</td>
</tr>
<tr>
<td>UK Home Office Research Database</td>
<td>20</td>
<td>135</td>
</tr>
</tbody>
</table>

Allard Feddes & Marcello Gallucci: A Literature Review on Methodology used in Evaluating Effects of Preventive and De-radicalisation Interventions
Table 2. Country in which the evaluation study of a counter-radicalisation intervention was conducted, number of samples and percentage of total number of samples

<table>
<thead>
<tr>
<th>Country</th>
<th>Number of samples</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algeria</td>
<td>5</td>
<td>4 %</td>
</tr>
<tr>
<td>Australia</td>
<td>1</td>
<td>1 %</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>4</td>
<td>3 %</td>
</tr>
<tr>
<td>Denmark</td>
<td>8</td>
<td>6 %</td>
</tr>
<tr>
<td>Egypt</td>
<td>2</td>
<td>1 %</td>
</tr>
<tr>
<td>Finland</td>
<td>1</td>
<td>1 %</td>
</tr>
<tr>
<td>Germany</td>
<td>8</td>
<td>6 %</td>
</tr>
<tr>
<td>Indonesia</td>
<td>4</td>
<td>3 %</td>
</tr>
<tr>
<td>Iraq</td>
<td>1</td>
<td>1 %</td>
</tr>
<tr>
<td>Israel</td>
<td>2</td>
<td>1 %</td>
</tr>
<tr>
<td>Jordan</td>
<td>2</td>
<td>1 %</td>
</tr>
<tr>
<td>Morocco</td>
<td>2</td>
<td>1 %</td>
</tr>
<tr>
<td>Netherlands</td>
<td>36</td>
<td>26 %</td>
</tr>
<tr>
<td>Norway</td>
<td>4</td>
<td>3 %</td>
</tr>
<tr>
<td>Philippines</td>
<td>2</td>
<td>1 %</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>9</td>
<td>7 %</td>
</tr>
<tr>
<td>Country</td>
<td>Count</td>
<td>Percentage</td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>-------</td>
<td>------------</td>
</tr>
<tr>
<td>Sweden</td>
<td>1</td>
<td>1%</td>
</tr>
<tr>
<td>Sri Lanka</td>
<td>1</td>
<td>1%</td>
</tr>
<tr>
<td>Thailand</td>
<td>1</td>
<td>1%</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>16</td>
<td>12%</td>
</tr>
<tr>
<td>United States of America</td>
<td>4</td>
<td>3%</td>
</tr>
<tr>
<td>Yemen</td>
<td>4</td>
<td>3%</td>
</tr>
<tr>
<td>Mixed number of countries</td>
<td>13</td>
<td>13%</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>135</td>
<td>100%</td>
</tr>
</tbody>
</table>
Table 3. Number of Samples (N = 135) Divided by Intervention Sample and Sample Ideology

<table>
<thead>
<tr>
<th>Sample</th>
<th>Individual</th>
<th>Group</th>
<th>Social context</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-radical</td>
<td>Potentially Radicalised</td>
<td>Non-radical</td>
</tr>
<tr>
<td>Islamic extr.</td>
<td>5</td>
<td>7</td>
<td>29</td>
</tr>
<tr>
<td>Extremism</td>
<td>6</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>RWE</td>
<td>2</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>Terrorism</td>
<td>-</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Separatist</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Ethn. camp.</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>13</strong></td>
<td><strong>23</strong></td>
<td><strong>50</strong></td>
</tr>
</tbody>
</table>

Note. Islamic extr. = Islamic extremism; Extremism = Extremism in general; RWE = Right-wing extremism; Terrorism = Terrorism in general; Separatist = National separatist; Ethn. camp. = Ethnic campaigning for compatriots abroad
Table 4. Number of Samples (N = 135) Divided by Intervention Sample and Intervention Goal

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Individual</th>
<th>Group</th>
<th>Social context</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-radical</td>
<td>Potentially radical</td>
<td>Radicalised</td>
</tr>
<tr>
<td>ST. prev.</td>
<td>2</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>ST. rest.</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ST. supp.</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LT. prev.</td>
<td>62</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>LT. rest.</td>
<td>49</td>
<td>5</td>
<td>37</td>
</tr>
<tr>
<td>LT. supp.</td>
<td>15</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>ST. prev. &amp; LT. prev.</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Allard Feddes & Marcello Gallucci: A Literature Review on Methodology used in Evaluating Effects of Preventive and De-radicalisation Interventions
<table>
<thead>
<tr>
<th></th>
<th>LT. prev. &amp; ST.</th>
<th>LT. rest. &amp; LT.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>rest.</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>supp.</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

### Table 5. Number of Samples Divided by Intervention Sample and Evaluation Focus

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Individual</th>
<th>Group</th>
<th>Social context</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-radical</td>
<td>Potentially radical</td>
<td>Radicalised Non-radical</td>
</tr>
<tr>
<td>Impact</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mechanism</td>
<td>1</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Process</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Economic</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Imp. &amp; Mech.</td>
<td>6</td>
<td>9</td>
<td>35</td>
</tr>
<tr>
<td>Imp. &amp; Proc.</td>
<td>2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Imp. &amp; Ec.</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Imp., Mech. &amp;</td>
<td>4</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Imp., Mech.,</td>
<td>Proc.</td>
<td>Econ.</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>Proc.</td>
<td>-</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Imp., Mech.,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proc. &amp; Econ.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>13</td>
<td>23</td>
<td>50</td>
</tr>
</tbody>
</table>