Exploring the transferability and applicability of gang evaluation methodologies to counter-violent radicalisation

Matthew Davies, Richard Warnes, Joanna Hofman
This report derives from one aspect of the research conducted under the IMPACT Europe project.\textsuperscript{1} IMPACT Europe is a project funded by the Seventh Framework Programme\textsuperscript{2} that seeks to identify what works in tackling violent radicalisation, or countering violent extremism (CVE), as it is often more recently referred to. A major part of the project was focused on developing an interactive evaluation toolkit which draws on current evaluations of CVE. Given that the main challenge relates to the limited number of high-quality evaluations of CVE interventions, examples of CVE evaluations collated by IMPACT Europe were insufficient to populate the evaluation toolkit and to advance the evaluation culture in the field. Consequently, given the need for usable and informative comparative research, we looked more widely at the more developed and rigorous evaluation studies in various areas of criminology.

\textsuperscript{1} For more information, see the IMPACT Europe website. As of 07/07/2017: http://impacteurope.eu/
\textsuperscript{2} For more information see the Programme website. As of 07/07/2017: https://ec.europa.eu/research/fp7/index_en.cfm
# Table of contents

Preface ........................................................................................................................................ iii
Table of contents .................................................................................................................. v
Figures and tables .............................................................................................................. vi
Summary .............................................................................................................................. vii
Acknowledgements ......................................................................................................... viii
1. Introduction ................................................................................................................. 1
2. Similarities and differences between gang involvement and radicalisation ............... 3
3. Method statement ....................................................................................................... 7
4. Results ......................................................................................................................... 11
5. Limitations, discussion and conclusion ....................................................................... 17
References ......................................................................................................................... 21
Figures and tables

Figures

Figure 1. Quality of evaluations in CVE and gang desistance literature .................................................. 12
Figure 2. Selected evaluation methods and approaches used in CVE and gang desistance literature ........ 13

Tables

Table 1. Total number of hits and relevant articles found per keyword in the academic databases ........... 8
Table 2. Examples of high-quality applicable evaluations........................................................................ 15
Summary

Background
There are currently few high-quality and publically available evaluations of interventions designed to counter violent extremism. Various initiatives, including the IMPACT Europe project, aim to introduce evaluation to this policy area where robust and rigorous evaluations have not yet become the norm.

Aim
The aim of this report is to illustrate how exploration of evaluations from established academic fields can be used to inform the development of evaluation practice in relatively novel research areas. Specifically, this report highlights a lack of empirical evidence about effective interventions to counter violent radicalisation and argues that there are relevant lessons on evaluation that can be drawn from studies of gangs. Through a comparative analysis of evaluations of gang interventions, this report examines the transferable and applicable lessons for conducting evaluations of CVE interventions.

Method
A targeted literature review was conducted to identify relevant papers in the field of criminology. In total, 66 papers were reviewed in full and information on background, intervention type, evaluation methods and quality was coded and extracted for descriptive analysis.

Results
A wide range of evaluation methods are used in gang literature, with experimental and quasi-experimental designs accounting for almost half of all evaluations. Impact evaluations are the most commonly used type of evaluation for gang desistence interventions. Evaluations of gang interventions were rated as being of higher quality than evaluations of CVE interventions, and almost all gang evaluations reviewed were either applicable and/or transferable to CVE.

Conclusions
The results illustrate the value of comparative research in developing evaluation approaches in the field of CVE by borrowing from gang-related interventions. It also demonstrates that in some areas, CVE research might offer useful insights to future gang evaluations.
Acknowledgements

The research team would like to acknowledge Dr Allard Feddes and Prof Marcello Gallucci of UNIMIB who supported the work. We would also like to thank Dr Megan Sim and Kate Cox at RAND Europe, who supported the data collection exercise. Our thanks also go to our quality assurance reviewers, Dr Alex Sutherland and Dr Tom Ling, for their insightful comments and suggestions.

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

1.1. Purpose and method of the study

There are currently few high-quality evaluations of interventions designed to counter violent extremism (van Hemert et al., 2014). While much has been written about the wider issues of extremism, radicalisation and terrorism, most of the existing literature on CVE interventions and evaluations of their effectiveness is anecdotal and lacking in empirical rigour (Romaniuk, 2015; van Hemert et al., 2014). This report sets out to address the empirical gaps in evaluations of CVE interventions by exploring the possible applicability and transferability of related criminological literature – in particular evaluations of interventions to disrupt or minimise the harmful impacts of gangs.

This report documents the process of conducting a targeted literature review and a comparative analysis between the fields of CVE and gang-desistance literature. It presents the results of this exploratory exercise, highlighting the relative rigour of gang-related evaluations in comparison to CVE evaluations. Building on these findings, the paper discusses some of the pertinent lessons that emerge from the rich research tradition in the field of gangs. The discussion illustrates the value of comparative research in evaluation methods, particularly in nascent academic fields. To our knowledge, this is the first attempt to systematically compare gang desistance literature with literature on CVE.

1.2. Terminology used

There is no single definition of radicalisation. While only a minority of people who have been radicalised have gone on to engage in acts of terrorism, all of those who have engaged in terrorism have at some point been radicalised. This understanding is reflected in various definitions of radicalisation adopted by the European Union and its Member States. For example, the UK defines radicalisation as:

\[\text{the process by which a person comes to support terrorism and extremist ideologies associated with terrorist groups}\] (Home Office, 2016, paragraph 7).

The Dutch intelligence services defines it as:

\[\text{an increasing willingness to pursue and/or support fundamental changes in society, possibly by undemocratic means, which are in conflict with or could pose a threat to the democratic legal order}\] (General Intelligence and Security Service, 2011).

The term ‘radicalisation leading to extremist violence’ (European Commission, 2015) is used at the European level to acknowledge that individual rights and personal freedoms are protected by the European Human Right Convention and the Charter of Fundamental Rights of the European Union.
Specifically, the European Commission defines radicalisation as ‘embracing opinions, views and ideas which could lead to acts of terrorism’ (European Commission, 2005). This is distinct from extremism, which is associated with ‘active adoption of an ideology, intending to deliberately apply violence to remove a state’s structure and its elite’ (European Parliamentary Research Service, 2015). In this report, the word radicalisation always refers to radicalisation leading to violent extremism unless otherwise stated.

Likewise, there is no universally agreed definition of gangs. Although organised criminal networks may be conceived of as gangs, for the purposes of this work we focused on street gangs, which have received substantial research attention in the last 80 years (although where possible we draw lessons from key pieces of work on organised crime – see for example Varese, 2012). We adopted the definition agreed by the Eurogang network:³

[...] a street gang (or troublesome youth group corresponding to a street gang elsewhere) is any durable, street-oriented youth group whose involvement in illegal activity is part of its group identity (Weerman et al., 2009, p20).

While this definition has come under some criticism (e.g. Smith and Egan, 2014), it is sufficiently broad to cover a wide array of research in the field, which may provide valuable lessons for CVE evaluations.

1.3. Structure of the report

The report starts by outlining similarities and differences between gang involvement and radicalisation (Chapter 2) and presents the methodology used (Chapter 3). The results of the comparative analysis across the field of CVE and gang desistence literature are reported in full (Chapter 4) and discussed with a view to the limitations of the study (Chapter 5).

³ Eurogang is a thematic network for comparative and multi-method research on violent youth groups in Europe. For more information, as of 07/07/2017: https://www.umsl.edu/ccj/Eurogang/euroganghome.html
2. Similarities and differences between gang involvement and radicalisation

We know little about CVE interventions from published literature. However, as we will argue, the similarities between those at risk of recruitment into both gangs and radicalisation groups mean that there should be opportunities for knowledge from one field to be applied to the other. Consequently, if terrorism is contextualised as a form of politically or ideologically motivated crime (as opposed to a form of low-intensity warfare), researchers such as LaFree and Miller (2008) have argued that there are similarities between terrorism and other more conventional forms of crime. Building on this, LaFree and Dugan (2004) assert that because of conceptual similarities between crime and terrorism, there should be some level of transferability. They argue that many criminological research methods could be applicable to the study of terrorism, while the knowledge base of crime research might be used to inform research on terrorism.

There are some key parallels between gangs and radicalisation, pertaining to demographics, recruitment, group dynamics, behaviour and desistance. With regard to demographics, both those drawn to gangs and to violent extremism are typically marginalised young males (Decker and Pyrooz, 2015). As with gangs, terrorist networks rely on a ‘self-selection’ process to identify suitable and motivated individuals for membership. Often the initial process consists of identifying individuals with the psychological ‘predisposition’ and ideological/political motivation to become involved, before subjecting them to a vetting and screening process and tests to identify and confirm their commitment and reliability. While in the case of gangs such tests will likely comprise initiation such as the commission of criminal acts, often including the use of violence, terrorist networks test potential members based on their willingness to participate in and/or support terrorist activities (Warnes, unpublished).

There are also similarities in terms of group dynamics. Research on street gangs has indicated that there is rarely a clear sense of hierarchy, and that they often resemble an ‘informal-diffuse’ structure with loose affiliation to the unit (Morselli, 2009). Similarly, research on terrorist networks indicates that organisational structure is relatively flat, and that membership is transient (Sageman, 2008; see also Curry, 2010). Both groups have similar recruitment processes and thrive on a common enemy (Decker and Pyrooz, 2015). With regard to behaviour, violence is a common feature – often threatened rather than actualised (Decker and Pyrooz, 2015), and often used as a form of ‘self-help’ or an attempt to address perceived wrongs (Curry, 2010). Perhaps most significantly for this paper, a number of academics in the field of terrorism have suggested that criminological research may provide transferable lessons in relation to desistance, which could prove applicable to CVE (Horgan, 2009). As Bjorgo and Horgan argue:
Many of the factors and processes involved in leaving terrorist organisations, religious cults, racist groups and criminal youth groups appear strikingly similar in spite of great differences in ideological context, background and aspirations (Bjorgo and Horgan, 2009, p.5).

Criminology has developed a rich body of research related to desistance, built on life-course theory, which may have direct relevance for counter-radicalisation (Bottoms et al., 2004; Decker et al., 2014; LaFree and Miller, 2008; McNeill, 2012; Shapland and Bottoms, 2011).

There are, however, some important differences between gang involvement and radicalisation. One significant difference lies in motivations. Whilst ideology is a key feature of radicalisation, a belief system is not integral to the existence of gangs (Decker and Pyrooz, 2011). Rather, gang members are motivated to join gangs by many other reasons (see e.g. Sanchez-Jankowski, 1991; Densley, 2013), such as profit (Curry, 2010). There are also differences in terms of geographic scope. Research on street gangs in the US context suggests that they are characteristically clustered in local geographic units, while radical networks operate globally, facilitated by technology (Decker and Pyrooz, 2011).

While it is important to be aware of these differences, this paper argues that there is value in drawing on evaluations of gangs to inform CVE evaluations. From an evaluation perspective, both gangs and radicalised groups involve hard-to-reach populations, rendering access particularly challenging. However, that does not mean that covert or otherwise ‘secret’ groups cannot be studied empirically (see Chattoe and Hamill, 2005; Densley, 2013; Varese, 2012; Gambetta, 1996). Evaluators also have to contend with the inherently hidden nature of much of these groups’ activity, since both populations may engage in criminal activity that is not always reported. There are also a series of misconceptions about the definitions, scale and scope of both gangs and radicalised groups, which cloud the respective research fields. Research on gangs has developed over the course of 90 years (see Thrasher, 1927) and has developed approaches to deal with some of these obstacles. Decker and Pyrooz (2015) observe that there have been over 5,000 works on gangs dating back to the 1920s, and a 50-year history of empirical research in the field (compared to 319 studies on radicalisation dating back to the 1980s, identified by Christmann (2012)). As such, the evidence base as regards evaluations of gang desistance interventions is larger and more robust than its CVE equivalent, and several meta-analytical studies have been published in the field (Disley et al., 2011).

As a result of the contextual similarities, the research into empirically grounded gang desistance literature – particularly examples of evaluations – was conducted in order to identify lessons that could be applied and transferred to the field of CVE. Other criminology-related literature was also considered by the research team. This included literature on cults and right-wing extremist groups since both share commonalities with radicalisation in terms of joining, participating and leaving (see Bromley, 2006; Olsen, 2006; Bjorgo, 2009). However, the literature in these fields is still in its infancy and there have been comparatively few evaluations of targeted interventions in these areas – particularly robust, empirical

---

4 A further distinction can be made in relation to the concept of ‘glocalism’ – a term used to describe how a nation-state’s political legitimacy is shaped by local insecurity under the forces of globalisation (Gurbuz, 2016). Although there is limited research in this area, it is easier to see how glocalism may be of more relevance to radicalisation processes compared to street gang formation.
exploring the transferability and applicability of gang evaluation methodologies to counter-violent radicalisation

valuations (Disley et al., 2011). As such, a central focus of this article is on the well-developed body of gang-related evaluations and their relevance for CVE evaluations.
3. Method statement

3.1. Literature search

The literature search took the form of a targeted review. Rather than providing a comprehensive review of literature, the aim was to identify a series of key papers in the field of criminology to illustrate where and how knowledge could be applicable and/or transferable to CVE evaluations.

We initially consulted eight crime-related papers that had been referenced by van Hemert et al. (2014). Seven of these articles matched the inclusion criteria and were added to the review. A relevant Rapid Evidence Assessment (REA) that was known to the team was also included in the review (Disley et al., 2011), and by searching the references in this REA, the team were able to identify a further seven relevant articles that were included.

Subsequently, peer-reviewed evaluations were systematically searched for through academic databases containing high volumes of criminology-related literature. The databases searched were: Web of Science, the UK Home Office Research Database, Science Direct and the Cochrane Library. We also searched Google Scholar.

The searches were conducted using a set of keywords to reflect the focus on evaluations of gang interventions. The search terms were as follows: Gang AND intervention AND evaluation OR assessment OR programme OR effectiveness.

The procedure for identifying relevant 'hits' began with browsing through journal titles in academic databases and opening abstracts where titles indicated a study with a gang intervention. An inclusive approach was used when browsing through titles, which entailed opening abstracts of studies even where titles were ambiguous and where it was unclear whether the study contained an evaluation of an intervention. This approach was used in order to avoid excluding potentially relevant studies at this stage. For the databases that had returned a high volume of hits (notably Science Direct and Google Scholar), only the first five pages were reviewed. Since the pages were ordered by relevance, it is unlikely that there would be relevant results beyond these pages (an approach consistent with van Hemert et al., 2014), although we acknowledge the limitations of this approach. The research team initially conducted preliminary checks of subsequent pages, which supported the validity of this choice.

---

5 We did not include the US spelling in searches of ‘programme’, although the search still identified a high volume of US studies.
3.2. Inclusion criteria

We included articles that were in English and had been published between 01/01/1995 and 31/12/2015. While a literature review on CVE between 2000 and 2014 was carried out by van Hemert at al. (2014), we extended this search period for gang desistance literature, since this literature has a longer and richer history of high-quality empirical evaluations. The total number of hits and relevant articles found in each academic database are presented by search term in Table 1 below. This search resulted in a total of 60 relevant articles (i.e. published results containing relevant keywords). However, following further detailed review of the articles, 22 were ultimately excluded from the review because they did not satisfy the inclusion criteria.

Table 1. Total number of hits and relevant articles found per keyword in the academic databases

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Web of Science</td>
<td>14 (30)</td>
<td>5 (20)</td>
<td>4 (84)</td>
<td>4 (18)</td>
</tr>
<tr>
<td>UK Home Office Research Database</td>
<td>0 (24)</td>
<td>0 (29)</td>
<td>0 (23)</td>
<td>0 (31)</td>
</tr>
<tr>
<td>Science Direct a</td>
<td>2 (2,542)</td>
<td>0 (3,224)</td>
<td>3 (4,235)</td>
<td>1 (2,048)</td>
</tr>
<tr>
<td>Cochrane Library</td>
<td>0 (2)</td>
<td>0 (1)</td>
<td>0 (2)</td>
<td>0 (2)</td>
</tr>
<tr>
<td>Google Scholar b</td>
<td>11 (41,500)</td>
<td>5 (45,500)</td>
<td>6 (23,700)</td>
<td>5 (31,200)</td>
</tr>
<tr>
<td>Total</td>
<td>27 (44,098)</td>
<td>10 (48,774)</td>
<td>13 (28,044)</td>
<td>10 (33,299)</td>
</tr>
</tbody>
</table>

Note: ‘Relevant’ articles were only counted when they were new (i.e. had not already been identified through other academic databases or search terms) to avoid recording duplicate articles.

a Only the first five pages of hits were reviewed.
b Only the first five pages of hits were reviewed.

To supplement these articles, we consulted additional gang-related resources. These included the Center for Program Evaluation and Performance Measurement in the US, which lists a large number of both

---

6 Additionally, several high-profile and robust gang evaluations took place in the US in the late 1990s, related to the Gang Resistance Education and Training (GREAT) programme, which the research team deemed important to include (see Esbensen et al., 2013).

7 Relevant articles here are those that appeared to match the inclusion criteria upon review of abstracts, rather than those that were eventually included in the gang database.
peer-reviewed journal articles and US government evaluations of gang interventions. In order to capture literature that was specific to the European context, we also consulted material from Eurogang, which helped to identify additional resources that had not been collected through academic databases. Twenty-four publications were discovered through this route, and although these were not peer-reviewed articles, the majority were large-scale national evaluations using mixed-method designs and were therefore considered relevant for inclusion in the review. In total, 66 articles were included in the review.

3.3. Definitions

Having selected articles for review, we then read through the introduction and methodology sections of all papers to determine whether they were of relevance. We identified relevant articles based on whether the study consisted of both an intervention that targeted gangs (either individual gang membership or broader gang-related crime) and an evaluation (including a broad range of study designs) of the intervention. In addition, we sought to identify evaluations that could be applicable and/or transferable to CVE. Applicability was defined as the usefulness of the evaluation design and methods (i.e. does the evaluation contain any elements that could add value to a CVE evaluation?). Transferability refers to the feasibility of transferring the evaluation design and methods (i.e. could the evaluation be practically implemented in a CVE setting?). These terms were included in order to clearly elicit key lessons for conducting evaluations of CVE interventions.

Upon identification of relevant articles, we devised a framework to code relevant variables from the literature. Using the framework from van Hemert at al. (2014), the variables were clustered under three overarching headings: ‘Background’ (where details of the publication were recorded), ‘Intervention Type’, and ‘Evaluation’. This approach was based on concepts similar to the ‘Truth Tables’ found in Qualitative Comparative Analysis (Ragin, 1989) and contained additional items such as ‘Group of Focus’, ‘Intervention Goal’ and ‘Evaluation Approach’, each of which was broken down further into further sub-variables to allow coding. For example, under the ‘Evaluation’ heading, ‘Evaluation Focus’ was a high-level variable which was divided into four options: ‘Economic’, ‘Impact’, ‘Mechanism’ and ‘Process’ (based on van Hemert at al., 2014).

A quality assessment of evaluations was also conducted based on scientific criteria taking into account the practical limitations that evaluators face when conducting field research (e.g. absence of a control group due to ethical constraints). The methods and instruments commonly used in evaluation research in the area of behavioural sciences were taken as a reference point. In evaluation research, and as outlined by Leeuw (2012), empirically testable assumptions and hypotheses should be outlined and tested and this is considered good practice. As such, the quality of evaluations was scored as follows:

---

8 As of 15/10/2015: https://ojp.gov/programs/gangs.htm
9 Although the gang database includes 66 distinct data entries, it contains 62 unique articles since one article was divided into five parts to account for discrete chapters that contained different types of intervention (Lafontaine et al., 2005). For ease of reference, these data entries are synonymously referred to as ‘articles’ throughout this section.
• Low – when no empirical investigation was conducted while circumstances would allow for a more thorough methodological assessment to answer key evaluation questions (assuming availability of sufficient financial and human resources);
• Medium – when empirical data were collected but the circumstances would allow for a more advanced data collection;
• High – when empirical data were collected using a multi-method approach comprising multiple instruments.

This approach was consistent with the one used by van Hemert et al. (2014) and although the bar was set very low and was somewhat subjective, it allowed us to compare the quality of CVE evaluations with those in the field of gang desistance.

The research team used a binary code to identify whether a variable was present (or not) in each of the 66 articles examined. This allowed the team to compare between publications and across specific variables. Articles were coded independently by three researchers. Ambiguous cases were discussed amongst the research team, and a fourth member of the team dip-sampled the coding to ensure consistency across researchers doing the rating. Descriptive analysis was then conducted in order to identify key emerging themes.
4. Results

The majority (77%, 51) of relevant articles that were identified entailed one evaluation of one intervention. Excluding the meta-analyses, commentaries and reviews, four articles involved multiple interventions with a single evaluation, while one study involved multiple evaluations of one intervention. Three articles had no specific intervention and evaluation component, but were included in the review as they were considered to have applicable and/or transferable material for evaluations of CVE interventions. As such, excluding the meta-analyses, commentaries and reviews, the review covers a total of 75 interventions and 67 evaluations.10

The initial analysis of the included articles revealed a number of key findings, which may be relevant to evaluations in the field of CVE.

Evaluations of gang interventions were rated as being of higher quality than evaluations of CVE interventions

We found that almost a third (30%, 20) of evaluations of gang interventions were of high quality, 55% (37) were of medium quality, and 15% (10) were of low quality. This highlights the rigour of gang evaluations in comparison to evaluations of CVE, where no high-quality evaluations were identified by van Hemert at al. 2014 (see Figure 1).

---

10 This is fewer than the 126 CVE evaluations that had been identified by van Hemert et al. (2014), although it should be noted that the majority of these (83%) were not peer reviewed.
Figure 1. Quality of evaluations in CVE and gang desistance literature

Experimental and quasi-experimental designs account for almost half of all gang evaluations

The research on gang desistance literature shows that around 47% (31) or nearly half of all gang evaluations involved quasi-experimental designs, and a further 14% (9) used experimental designs (Figure 2). Longitudinal research was also used in a third (33%, 22) of studies, while cross-sectional approaches also featured frequently and were identified in 27% (18) of articles. It should be noted that these categories were not mutually exclusive and some studies used a combination of approaches (see for example McGarrell et al., 2013). In comparison, only 2% (3) of CVE evaluations were quasi-experimental, 3% (4) were longitudinal, while just 1% (1) used both longitudinal and quasi-experimental designs (van Hemert et al., 2014). These findings in the area of gang desistance demonstrate a useful learning opportunity for applying academic rigour and empirical evaluation in the CVE field.
Over three quarters (76%, 51) of gang desistance evaluations last longer than six months, compared to 5% (6) in the research previously conducted on CVE interventions and their evaluation (van Hemert et al., 2014). Nearly all (93%, 117) evaluations of CVE interventions were cross-sectional studies, meaning they are limited to a ‘snapshot’ in time.

A wide range of evaluation methods are used in gang literature

Similar to evaluation approaches, the specific evaluation methods used in the literature on gangs were diverse, and 39% (26) used more than one evaluation method. Of all evaluations, 38% (25) reported using qualitative interviews, while just over a third (34%, 23) used surveys. Data mining (33%, 22) and observations (30%, 20) were also reported to be used in approximately a third of all cases.

Impact evaluations are the most commonly used type of evaluations for gang desistence interventions

Almost two thirds (63%, 42) of gang evaluations focus on their impact, while 21% (14) evaluate both process and impact. In comparison, almost half (49%, 62) of the CVE studies were impact and mechanism evaluations (van Hemert et al., 2014).

The majority of evaluations are conducted independently

The gang literature review shows that out of the 60 evaluations where the independence of the evaluator (i.e. whether they were conducted by an external evaluator) could be determined, 96% (58) were
Almost all gangs evaluations reviewed are applicable and/or transferable to CVE

One of the key questions relating to the work on gang desistance literature was the level of ‘applicability’ and ‘transferability’ of any evaluations of interventions to the associated field of CVE. This further research found that out of the total of 66 gang desistance articles reviewed, 95% (63) were classed as ‘transferable’ and/or ‘applicable’ to the CVE field. Exactly half (33) were coded as both applicable and transferable, a third (22) were coded as only applicable (and not transferable), and no evaluations were coded as transferable and not applicable. Overall, 83% (55) of the sources were classed as applicable to evaluations of CVE interventions, while half were considered transferable (i.e. could be practically implemented in a CVE setting).

Examples of evaluations that were coded as applicable but not transferable are presented in Table 2 below. Additional analysis was conducted on the 17 evaluations that were coded as high-quality. All the high-quality evaluations were deemed to provide elements that could be usefully applied to the field of CVE, while 59% were considered to be transferable only (and not applicable).
Table 2. Examples of high-quality applicable evaluations

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Applicability</th>
<th>Transferability</th>
</tr>
</thead>
<tbody>
<tr>
<td>YouthARTS Development Project (Clawson and Coolbaugh, 2001)</td>
<td>The national evaluation of the YouthARTS Development Project has aspects that can be applied in a CVE setting. For instance, the evaluation used a longitudinal design, collecting data at multiple time points at the start of, during and after the intervention, which can be applied to a CVE evaluation. The evaluation also used a quasi-experimental approach to compare the outcomes for programme participants with those of non-participants; this approach can be applied to a CVE evaluation. In addition, the evaluation employed methods such as surveys, skills assessments, focus group interviews and collection of administrative data – all of which can be applied in a CVE setting.</td>
<td>While it may be feasible to transfer the design (a longitudinal quasi-experimental design) of the YouthARTS evaluation to a CVE setting, the methods used (i.e. art knowledge and attitude and behaviour surveys, skills assessments, focus group interviews and collection of administrative academic and court data) might have to be adjusted for a CVE evaluation. For instance, while the attitudes and behaviour survey might be transferable, it is unlikely that the art knowledge survey would be relevant for a CVE evaluation (unless it was also an arts programme).</td>
</tr>
<tr>
<td>An impact evaluation of juvenile probation projects in Christian, Peoria, and Winnebago Counties (Wassenberg et al., 2002)</td>
<td>The evaluations of juvenile probation projects in Christian, Peoria and Winnebago Counties used a variety of methods, including field observations, document analysis, interviews, focus groups and monitoring and police data. These methods can be applied in a CVE setting.</td>
<td>While the methods used in the evaluations of juvenile probation projects in Christian, Peoria and Winnebago Counties are applicable in a CVE setting, they are not directly transferable. For instance, the stakeholders chosen for interview – including judges, prosecutors and probation officers – may not be relevant in a CVE evaluation. Additionally, police data on criminal arrests may not capture radicalised or violent behaviour.</td>
</tr>
<tr>
<td>A systematic meta-review of evaluations of youth violence prevention programs (Matjasko et al., 2012)</td>
<td>Cross-sectional evaluation may be applicable to a CVE setting when identifying factors that are significantly related to reducing violence and related behaviours.</td>
<td>Systematic reviews and meta-analysis may not currently be transferable to a CVE setting, given the limited volume of existing evaluations on CVE interventions.¹¹</td>
</tr>
</tbody>
</table>

¹¹ Although it is recognised that null results are informative (hence ‘applicable’), for the purposes of analysis conducted here, meta-analyses were not considered to be currently ‘transferable’ to the few evaluations that have been conducted on CVE to date. For a discussion of ‘null reviews’ see (as of 07/07/2017): https://www.spi.ox.ac.uk/research/details/how-empty-are-empty-reviews.html
5. Limitations, discussion and conclusion

5.1. Limitations of the study

One limitation of this study is that the review includes some overlap between articles that have been included in the gang review and those that were included as part of meta-analyses, commentaries and reviews. In total, four meta-analyses and three commentaries or reviews were included in the gang review. Each of these meta-analyses, commentaries and reviews considered multiple interventions and evaluations relating to gangs, many of which were already included in our review. Although this resulted in repetition of some studies (i.e. the inclusion of evaluations as individual sources and as part of another source), it was nevertheless important to include the most robust studies as individual cases in the review in order to elicit good evaluation practice and key lessons that could be applied and transferred to CVE.

A further limitation is that the majority of studies (75%, 50) of gang interventions were conducted in the US. One reason for this disproportionality may be that some European governments do not acknowledge having a gang problem, which means that no specific interventions have been designed, conducted and evaluated within these countries (Weerman et al., 2009). But even where there may be a perceived gang problem, the response may not be to intervene through specific programmes as has been the case in the US; instead governments (in particular those characterised by a welfare model) may respond at a more holistic level through investments in social institutions such as education, employment and health (Carlsson and Decker, 2005).

This is not to say that there have not been any gang-specific interventions in Europe, but this helps to explain why there have been fewer evaluations in a European setting. Nevertheless, the specific evaluation methods and approaches adopted in the US-based interventions were considered to be transferable to the European context. Moreover, it is not necessarily the number of papers reviewed that determines the utility of such a review, but the scope and quality of the lessons that can be drawn from the source material. Even one highly informative hit on a topic may be more useful as a reference than a series of low-quality hits.

Finally, it should be noted that the research focused on publicly available evaluations that were published between 1995 and 2015. As such, this review does not include any unpublished government evaluations.
5.2. Discussion

This study aimed to explore the feasibility of applying and transferring evaluation methods from the established body of research on gangs to a CVE context. The results illustrate the greater rigour of evaluations of gang interventions compared to those in the CVE field. The gang literature tends to offer higher-quality evaluations, is more likely to focus solely on impact and to use quasi-experimental designs, and typically covers longer periods of time. This is unsurprising in light of the fact that criminology has developed a strong body of research into gangs over the last fifty years compared to CVE.

A greater proportion of evaluations were considered to be applicable than transferable. One of the reasons for this discrepancy is that although many evaluations were perceived as potentially useful in a CVE setting (‘applicable’), it was more challenging to envisage how particular designs and methods could feasibly be used in evaluations of CVE interventions (‘transferable’). For example, several impact evaluations of gang desistence programmes drew on police records to indicate whether an intervention affected gang behaviour. However, gang behaviour is associated with a range of criminal activities, such as drug-dealing, violence or homicide – measures that can be readily captured with criminal justice statistics (e.g. police-recorded data). However, violent extremism may not manifest itself in criminal behaviour to the same degree and therefore is not often likely to be captured by police-recorded data. Consequently, while the use of criminal statistics may be applicable to evaluations of some forms of violent extremism, it may not always be feasible as an outcome measure in CVE interventions.

Overall, the high level of applicability reinforces both the validity of examining gang desistance literature for examples of interventions that can be applied to the less academically mature area of CVE, and the importance of conducting detailed evaluations of those interventions.

From this exploration, and building upon the work of Decker and Pyrooz (2015), some key methodological and content-specific lessons emerge. Firstly, Decker and Pyrooz (2015) emphasise the importance of attaining a balance between theory generation and theory testing. They observe that much gang research began in theory generation, with little proportionate empirical support. However, in last 20 years there has been a shift towards more empirical approaches to testing theories. Literature on extremism appears to be in a similar place to that of early gang research (and criminology more generally – see Wikström (2011)), characterised by an over-abundance of theory and lack of testing (e.g. Borum, 2011a; Borum, 2011b; McCauley and Moscalenko, 2008; Moghaddam, 2007). In the case of radicalisation, this may be due to the growing attention on the issue, as well as the comparative rarity of terrorist events and radicalisation processes (Gambetta and Hertog, 2016).

Secondly, definitions need to be carefully considered as part of the study design, since these will shape the unit of study. For example, Maxson and Klein’s (1990) study of gang homicides illustrated that broadening the definition of homicide doubled the number of police-recorded homicides in their study.\footnote{The definition was broadened from homicides that ‘further the collective goals of the gang’, to simply homicides that involved a gang member.} Likewise, careful distinctions are needed between terms such as radicalisation, extremism, and terrorism (Decker and Pyrooz, 2015). Particularly pertinent to the case of radicalisation, there are important questions about whether to define radicalisation by someone’s beliefs or their actions. Gang research has
demonstrated the value of distinguishing between *instrumental* and *symbolic* activities (Klein and Maxson, 2006). Gangs use more rhetorical (symbolic) violence than actual violence (e.g. Decker and van Winkle, 1996). Social media provides another outlet for symbolic violence, an issue similarly identified in relation to radicalisation (Conway, 2010). It is therefore important to conceptually separate symbolic from actual behaviour in terms of assessing outcomes of interventions.

Related to this, it is also essential to consider using methods that are shaped by a clearly defined unit of analysis. In the case of gangs, too often research has focused at an individual rather than a group level (Decker and Pyrooz, 2015). However, research on organised crime has provided a broader focus on social opportunity structures that facilitate or inhibit involvement in criminal groups (e.g. Kleemans and de Poot, 2008). In particular, social network analysis has been used to identify the centrality of key members of criminal networks and their relative influence on criminal activity (Campana and Varese, 2013; Campana, 2016). Adapting similar methods with a focus at the group level could shed further light on how gangs and radical networks operate and, crucially, on processes for desistance.

Finally, it is evident that triangulation forms the basis of robust, high-quality evaluations as it seeks convergence, corroboration and correspondence of results from the different methods (Greene et al., 1989). This is particularly important in the fields of gangs and radicalisation, characterised by hard-to-reach populations, where there are high risks of bias and limitations in collecting and analysing data. For example, under-reporting of crime (and any other activity associated with radicalisation) creates ‘dark figures’, making it essential to use a mixed-methods approach to piece together missing official data. The most promising evaluations of gangs use a range of official law enforcement data (e.g. arrests, seizures, assault locations) alongside primary data collection (e.g. surveys and interviews) to assess the effect of interventions on membership (e.g. Wassenberg et al., 2002). Although some may have reservations about self-report in relation to criminal acts, there is a body of research showing the validity of this approach (e.g. Farrington, 1999).

Although gang evaluations are currently more robust than CVE evaluations, they have notably neglected some areas of enquiry, which also have relevance to radicalisation. Future CVE evaluations ought to avoid such mistakes – indeed, in some areas CVE research may already have some advantage in this respect.

Firstly, Decker and Pyrooz (2015) caution against overlooking the role of women. There has been very little research on female gangs, but more broadly there has been little examination of the role of gender in moderating or exacerbating conflict (for an overview in the context of gangs, see Panfil and Peterson (2015)). This links to the point about neglecting to study gangs in their entirety (rather than at an individual level). CVE evaluations going forward could benefit from broadening the scope of research and capturing the role of gender (and other demographics) in the (de)radicalisation process. Given the attention this area of research has recently received (see Chowdhury Fink et al., 2013, Couture 2014, RAN 2015), one can expect that evaluations would follow suit.

Secondly, prisons can play an important role for both gangs and radicalisation. Prison can re-affirm roles and beliefs and create new cliques, but little work has been done in this area with gangs outside of the US context (e.g. Skarbek, 2014), related to research challenges such as access to subjects. There has already been some work examining the role of prisons in radicalisation (e.g. Veldhuis et al., 2014; Disley et al.,
2011; Istiqomah, 2011; Chowdhury Fink and El-Said, 2011), and this is an important step (one which could result in lessons for gangs research).

On a related point, the high proportion of research emanating from US illustrates how studies are skewed towards specific contexts. Work elsewhere has highlighted the absence of experimental or quasi-experimental studies of preventive gang interventions in low- and middle-income countries – areas where gang violence may have particularly high social and economic costs (Higgins et al., 2015). This challenge is not unique to gangs and speaks to a broader critique of the dominance of North American and European voices across the social sciences (see for example Connell (2007)).

Lastly, it should be noted that while radicalisation literature remains in its infancy, it is a fast-growing field and has already begun to create an informative knowledge base. For example, there are questions about the role of technology in facilitating group-based violence – an issue that has been explored already in the case of CVE (e.g. Conway, 2010; von Behr et al., 2013; Conway, 2016). In turn, this kind of knowledge may provide gang research with valuable lessons in the future.

Overall, this research highlights how systematically mapping out the contours of a well-established research area can bring value to nascent fields of study. This concept is of course not new: there is a substantial body of literature that explores whether, how, and when cross-fertilisation across different policy areas can occur (e.g. Graham et al., 2013; Meseguer and Gilardi, 2009; Stone, 1999). Some authors have argued that more attention should be paid to a more systematic approach to tackling the questions of when and how diffusion takes place (Graham et al., 2013). Focusing on one of the policy stages, namely evaluation, this report provides a comprehensive assessment of whether and how such learning and cross-fertilisation in evaluation practice can take place between two different yet comparable fields. As such, the report contributes to a wider debate on methodological development and innovation in evaluation (e.g. Picciotto, 2012; Greene and McClintock, 1991).

### 5.3. Concluding statement

This analysis has demonstrated the value of comparative research in developing evaluation approaches in the field of CVE, borrowing from gang-related interventions. It has also shown that in some areas, CVE research might offer useful insights into future gang evaluations which have so far largely overlooked the role of gender, imprisonment and technological development in the radicalisation process. Of course, the topics of gangs and CVE differ substantially and evaluation approaches will need to be tailored accordingly. However, comparison of methods and approaches is valuable in developing an overall evaluation framework that can be adapted. As part of the comparison, this analysis has demonstrated the value in distinguishing between applicability and transferability, which can help evaluators to devise rigorous evaluation designs. In particular, this research has demonstrated that there is value in comparative research in developing evaluation methodologies – not only with regard to borrowing and adapting evaluative approaches, but also in terms of learning from key lessons where other fields have fallen short.


Exploring the transferability and applicability of gang evaluation methodologies to counter-violent radicalisation


Exploring the transferability and applicability of gang evaluation methodologies to counter-violent radicalisation


2013. As of 07/07/2017:
http://www.rand.org/pubs/research_reports/RR453.html

