Through the Looking Glass
Assessing the Evidence Base for P/CVE Communications

Michael Jones
189 years of independent thinking on defence and security

The Royal United Services Institute (RUSI) is the world’s oldest and the UK’s leading defence and security think tank. Its mission is to inform, influence and enhance public debate on a safer and more stable world. RUSI is a research-led institute, producing independent, practical and innovative analysis to address today’s complex challenges.

Since its foundation in 1831, RUSI has relied on its members to support its activities. Together with revenue from research, publications and conferences, RUSI has sustained its political independence for 189 years.

The views expressed in this publication are those of the authors, and do not reflect the views of RUSI or any other institution.

Published in 2020 by the Royal United Services Institute for Defence and Security Studies.

This work is licensed under a Creative Commons Attribution – Non-Commercial – No-Derivatives 4.0 International Licence. For more information, see <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

RUSI Occasional Paper, July 2020. ISSN 2397-0286 (Online).

Royal United Services Institute
for Defence and Security Studies
Whitehall
London SW1A 2ET
United Kingdom
+44 (0)20 7747 2600
www.rusi.org
RUSI is a registered charity (No. 210639)
Contents

Introduction 1
  Methodology 2
  The Evidence Base 4

I. Limitations and Shortcomings in P/CVE Communications Approaches 9
  Ideological Preoccupations 9
  Scale, Resonance and ‘Message Dominance’ 17
  Credible Messengers and the Role of the State 25

II. What Looks More Promising? 31
  Improving Critical Thinking and Media Literacy 31
  The Right Messenger in Conversations and One-to-One Engagement 33
  Positive Alternatives Within Integrated Interventions 37
  Prioritising Process 41

Conclusions 45

About the Author 49

Annex I: Bibliography 51

Annex II: Research Methodology 65
Introduction

This paper focuses on P/CVE communications interventions as part of the Prevention Project, a broader appraisal of the available evidence base which examines ‘what can work and what has not worked’ in preventing and countering violent extremism (P/CVE). Funded by the Norwegian government, this research was started in January 2018 and aims to collate, analyse and strengthen knowledge in the prevention space.

Recognising that terrorism ‘is not simply violence but communication’,¹ P/CVE communications have become a prominent, if not staple, strand of preventive policy and programming. Designed to discredit, counter or confront extremist messaging, or strengthen the digital literacy and critical thinking capabilities of recipient societies, these measures include a diverse spectrum of interventions, both on- and offline.

Despite their ubiquity, this paper reveals significant shortfalls in the evidence base and emphasises the limitations and ambiguities of many communication-based activities. A paucity of empirical data on effectiveness, fragmented or outdated theoretical foundations and assumption-based logics constrain projects, particularly counternarratives, which attract disproportionate attention in the literature. These limitations are frequently accentuated by the insularity of messaging exercises, a preoccupation with ‘message dominance’ and a neglect for factors that make extremist content appealing in the first place. Crucially, the wider ecology framing how people consume and interpret content, and the dynamics of radicalisation and recruitment processes themselves, are often overlooked.

This review of the literature suggests there is anecdotal evidence that interventions are stronger when they recognise and respond to the social and relational dynamics of these processes. The use of credible messengers when engaged in dialogue and integrative approaches that leverage communications strategically to supplement messaging with wider activities appear promising. Programmes prioritising process and providing opportunities for narrative therapy, social change and individual empowerment are also encouraging as they strengthen inclusion and the agency of beneficiaries. However, these approaches continue to be underdeveloped in the literature and there is little analysis or experimentation into methods for expanding the scale and scope of such labour-intensive, highly tailored activities. Until these gaps are addressed, potentially effective interventions will struggle to create and sustain traction offline, or out-compete the ‘swarmcast’ online.²

The paper first briefly outlines the methodology and definitions used in this research. It then summarises the extent and limitations of the available evidence base before analysing the shortcomings and constraints of communication-based interventions, and identifying approaches that potentially show more promise. The final chapter summarises key conclusions and their associated recommendations.

**Methodology**

The paper reviewed 139 English-language studies that focused on P/CVE communications. These include peer-reviewed publications, independent evaluations, programme documents, and analytical and discursive grey literature (materials and research produced by organisations outside the traditional commercial or academic publishing and distribution channels). As outlined in the project methodology, each study was assessed according to quality (high, medium or low) on the basis of criteria drawn from the UK’s Department for International Development’s recommended practice, including the paper’s conceptual framing, transparency, methods, research design, validity, cogency and independence. Although the studies have been graded according to quality, the project authors have refrained from listing the grades of each reviewed paper in the publication series, given the potential biases and limitations detailed in Annex II.

The findings of each study were subsequently coded as ‘effective’, ‘potentially effective’, ‘mixed’, ‘ineffective’ or ‘inconclusive’, and analysed to identify limitations and more promising approaches in P/CVE communications.

It is important to note from the outset that the visibility of certain organisations, and the public availability of reports covering their work, inevitably means their interventions were more frequently cited and assessed in this review. This should not be seen as an attempt to undermine their work, and it is important to highlight the value of open access analysis in helping expand the evidence base. The project team also acknowledge that organisations and programmes are likely to have compelling information in support of their findings and theories of change that are not always accessible or included in publicly available documentation, often due to the data’s sensitive or confidential nature. However, the team was only able to analyse and draw conclusions from the material gathered through the review and have therefore tried to qualify findings on this basis.

In the context of P/CVE, communication activities include a spectrum of intervention types featuring ‘strategic communications’, ‘counternarratives’ and ‘alternative narratives’, with each category driven by discrete – albeit sometimes overlapping – objectives and mechanics.

---

3. For a full bibliography, see Annex I.
4. For more details on the project methodology, see Annex II.
Theoretically, these activities mutually reinforce each other when coordinated and managed effectively, delivering a holistic response to violent extremist (VE) content that helps mitigate the limitations of censorship and other ‘negative measures’. The literature generally coalesces around the taxonomy outlined below:

**Counternarratives**: Defined by Alejandro Beutel and colleagues as ‘a system of stories that hang together to provide a coherent view of the world for the explicit purposes of combating violent extremist narratives, and eliciting legal and non-violent activities in support of individuals, groups or movements, which support that worldview’. They are inherently reactive. As Michael Bamberg and Molly Andrews argue, counternarratives ‘only make sense in relation to something else’, thereby setting their positional and intentional characteristics. As such, they can be considered direct, often confrontational, efforts to deconstruct and demystify VE messaging, ‘whether through ideology, logic, fact or humour’.

**Alternative Narratives**: Designed as ‘positive stories’ to undercut extremist narratives by focusing on what we are for rather than against. They do not necessarily directly address VE content but may have ‘secondary outcomes’ that displace it by proactively advocating ‘anti-violent’ behaviour.

**Strategic Communications**: Considered as ‘an organisation’s use of the full range of communications channels to achieve a strategic objective’.

These terms are often conflated or used interchangeably in the literature. This is particularly the case with ‘counternarrative’, which has served as a vague and poorly defined label for anything from a ‘simple rebuttal’ to ‘programmes of strategic communication’. It also tends to be used synonymously with ‘counter-messaging’, a confusion between form and content.

---

11. *Ibid*.
that has important ‘real world’ implications.\textsuperscript{16} Despite their limitations, these categories are used to map P/CVE communications in this paper as they continue to be used by practitioners and evaluators.

**The Evidence Base**

‘P/CVE communications’ was the most populated thematic category in the Prevention Project, comprising 139 relevant studies. However, this aggregated figure is slightly misleading given that 65 studies focused on counternarratives specifically, with an additional 43 studies examining counternarratives within a wider package of initiatives. This suggests a rather reductive approach by implementers, as a potentially diverse suite of communications-based interventions is, in practice, overwhelmingly focused on a specific type of activity. Similarly, 60 of the studies covered digital interventions exclusively, with a further 37 addressing both on- and offline projects.

Despite the volume of material available, therefore, the corpus of literature collected in this review reveals a lack of breadth and scope. Studies generally allude to an expansive range of programming, but recycle the same small cluster of case studies when providing either programme results or a detailed analysis of an intervention’s implementation. Many of these examples can be traced back to a small London-based ‘lattice’ of NGOs and think tanks.\textsuperscript{17} This could be due to certain organisations displaying greater transparency in a context generally defined by a lack of data sharing. However, what amounts to a relatively small tranche of projects\textsuperscript{18} is repeatedly framed as exemplifying a competitive, crowded and highly innovative counter-messaging industry. This may contribute to a ‘disconnect between the way counter narratives are spoken about in policy papers or political circles and the actual quantitative scale they are operating on’, as there is little analysis about or evidence from interventions outside these case studies, making it unclear what they actually involve, their scope and how they have performed.\textsuperscript{19} Moreover, many frequently cited examples such as Average Mohammed or

\textsuperscript{16} Ibid.


\textsuperscript{18} Frequent case studies include ‘Abdullah X’, ‘Average Mohammed’, ‘Extreme Dialogue’, ‘Think Again Turn Away’ and ‘The Redirect Method’.

\textsuperscript{19} Digby, ‘Cartoons vs. the Caliphate’, p. 40. There are exceptions, for example, many counter-narrative and alternative-narrative projects are catalogued in guidance material such as Lilah Elsayed, Talal Faris and Sara Zeiger, ‘Undermining Violent Extremist Narratives in the Middle East and North Africa: A How-To Guide’, Hedayah, December 2017; RAN, ‘Preventing Radicalisation to Terrorism and Violent Extremism: Delivering Counter or Alternative Narratives’, RAN Collection of Approaches and Practices, 2019; and Vidhya Ramalingam, ‘On the Front Line: A Guide to Countering Far Right Extremism’, ISD. However, many cases studies tend to briefly outline activities without supplementary monitoring or evaluation data, or to focus almost exclusively at the output level.
Abdullah X – two online-based counter-messaging initiatives – were pilot schemes but are often repackaged in the literature as ‘long-running strategies’.  

Similarly, only 33 of the included studies were programmatic evaluations. The remaining 106 were policy reports, grey literature and academic studies that primarily focused on conceptual or theoretical framing. The field has also become increasingly saturated with guidance manuals, instructive content and counternarrative libraries to strengthen monitoring and evaluation regimes and recommend ‘good practice’. Consequently, a significant portion of the literature focuses on technical prescriptions rather than providing detailed analyses of interventions or interrogating the various assumptions that often underpin counter-messaging approaches.

The literature was also remarkably insulated. While there were exceptions, this review found few references to cross-cutting lessons or theory in studies focusing on P/CVE communications, despite the relevance of various disciplines including theology, psychology, psychosocial studies, criminology, conflict studies, advertising, mass marketing, cyber security, literary studies and communication studies.

Table 1: Summary of the Team’s Assessment of P/CVE Communications Interventions

<table>
<thead>
<tr>
<th>Intervention Impact</th>
<th>Quality of Evidence</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Moderate</td>
</tr>
<tr>
<td>Effective</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Potentially effective</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Mixed</td>
<td>12</td>
<td>24</td>
</tr>
<tr>
<td>Ineffective</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Inconclusive</td>
<td>16</td>
<td>29</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>41 (29%)</strong></td>
<td><strong>68 (49%)</strong></td>
</tr>
</tbody>
</table>

Source: Author generated.

In terms of the findings, the largest share of studies (48%) was considered ‘inconclusive’, followed by those categorised as ‘mixed’ (30%), as summarised in Table 1. Only 16% of studies, mostly of high or moderate quality, were found to show ‘potentially effective’ results, and none of the collected studies were assessed as presenting clearly ‘effective’ interventions. As discussed in this paper, the majority of potentially effective interventions were multifaceted programmes that included the circulation of counternarratives and alternative narratives as part of a wider range of activities.

Conversely, only 6% of studies clearly demonstrated the ‘ineffectiveness’ of communication-based interventions. This reflects a broader qualification of the review’s analysis and findings, as a lack of data should not necessarily be conflated with evidence of failure.

The high proportion of ‘inconclusive’ studies reveals a general paucity of data documenting the effectiveness of counternarratives or linking interventions with assessed outcomes. For example, of the nine example projects detailed in Charlie Winter and Johanna Fürst’s analysis of counter-speech practices across the UK, France and Germany, only one provided results beyond the output level and most did not include any success measure at all. In the collected literature, some studies list quantitative findings but these routinely succumb to the allure of ‘vanity metrics’ such as reach, views and engagement – a feature shared by many appraisals of online programming that gives little insight into cognitive, attitudinal or behavioural shifts, or the offline conduct of audiences. Empirical evidence at the impact level was sparse, as were longitudinal analyses tracking longer-term influence and sustainability. While advances have been made in social media analytics, efforts to systematically measure the influence of interventions remain subjective, abstract and largely unverified, often leaving assessments


24. ‘Vanity metrics’ is a marketing term for seemingly impressive but somewhat superficial measures that give little insight into the true performance or impact of a project.


27. Ibid.
dependent on process evaluations,\textsuperscript{28} anecdotal observation and inductive leaps.\textsuperscript{29} Consequently, ‘wishful strategies’ and ‘inconclusive’ findings are pervasive across the literature.\textsuperscript{30} These trends are shared with the wider P/CVE space, raising concerns of pre-packaged solutions applied to problems, irrespective of need, context or relevance.\textsuperscript{31}

This is not to diminish the difficulty of curating material on communication-based interventions in a systematic way nor determining its impact.\textsuperscript{32} The relevant data is often ephemeral, requiring ‘continuous collection, formatting, and archiving’,\textsuperscript{33} far beyond the capacities of typical P/CVE initiatives. Nevertheless, it makes monitoring and evaluation difficult and severely limits the strength of the available evidence base.

P/CVE communications also seem to be grounded in a set of implicit and explicit assumptions rather than a ‘fully articulated theory’, as discussed in the analysis below.\textsuperscript{34} These are widely acknowledged in the literature, and yet are not only recycled as received wisdom without any corroborating evidence but often continue to frame project design. This can perhaps be partially attributed to the urgency and uncertainty of the terrorist threat, particularly in the online space, encouraging a logic of ‘let’s just try something’ where policies and programmes are manufactured with little concept of ‘why and how they interact with other perceptions, policies or actors’\.\textsuperscript{35} However, it leaves communication-based interventions – and counternarratives in particular – dependent on ‘shaky theoretical and empirical foundations’ that appear to constrain their effectiveness in many cases.\textsuperscript{36}

\begin{itemize}
\item \textsuperscript{29} Henry Tuck and Tanya Silverman, ‘The Counter-Narrative Handbook’, ISD, 2016.
\item \textsuperscript{30} Berger and Strathearn, ‘Who Matters Online’.
\item \textsuperscript{34} Glazzard, ‘Losing the Plot’.
\item \textsuperscript{36} Glazzard, ‘Losing the Plot’, p. 1.
\end{itemize}
I. Limitations and Shortcomings in P/CVE Communications Approaches

Drawing on analysis of the evidence base, this chapter examines the limitations and shortcomings of P/CVE communications interventions, with many approaches often being constrained by preoccupations with countering ideology and achieving ‘message dominance’. This tends to neglect the factors which make extremist content appealing in the first place, the wider ecology of how people consume and interpret content, and the dynamics of recruitment and radicalisation processes themselves. Such problems are frequently accentuated by the insularity and siloing of online messaging campaigns and a reliance on problematic or unsuitable messengers.

Ideological Preoccupations

The literature on counter-messaging included in this review tends to focus on ‘battling’ terrorist ideologies, with projects contesting, refuting and contradicting VE narratives to undermine the ‘naturalising’, ‘obscuring’, ‘universalising’ and ‘structuring’ propensities of extremist claims. This frequently involved promoting a particular set of ideological positions or theological interpretations. However, there are theoretical questions over the role of ideology in radicalisation and recruitment, and in countering such processes. There is also limited empirical evidence of the effectiveness of ideological and religiously oriented interventions. Additional issues and problematic assumptions arise from the idea of moderation and the agency of ‘moderate actors’ in leading these counter-ideological efforts.

It should be acknowledged that ideology is not synonymous with religion: the former generally features some form of worldview, advances a model of a desired future and explains how change ‘can and should be brought about’. This can include – but is not necessarily dependent on – theological ideas, while many ideologies are avowedly secular. However, with P/CVE’s historical

focus on violent extremism in its ‘Islamist’ iteration, formal and informal interventions are often based on theological arguments to refute claims made by militant groups or push non-violent interpretations of scripture. Campaigns cited in the literature include many involving ecclesiastical authorities circulating rebuttals on social media, extremist chatrooms and YouTube. ‘Radical New Way’, for instance, runs a website devoted to capturing the attention of Anglophone audiences ‘similar to the one targeted by AQ [Al-Qa’ida], by posting Islamic opinions, teachings, and mediating discussions all directed toward non-violence’. The Muslim Public Affairs Council in the US subscribes to a similar ‘e-dawah’ approach, distributing ‘original video content’ such as ‘Injustice Cannot Defeat Injustice’ to publicise ‘popular American imams condemning violence in the name of Islam’. According to Todd C Helmus and colleagues, social media platforms can function as ‘virtual mosques’, with some sites like suhaibwebb.com drawing 70,000 hits per day and ‘rock-star imams’ using their credibility to amass vast social networks and act as ‘personal counsellors’ when ‘followers reach out to them directly via social media’. Anne Speckhard similarly cites the Salafi scholars of Siraat, who ‘monitor English militant jihadi sites’ and then ‘strategically post Salafi-based scholarly papers’ to successfully inject doubt about the veracity of terrorist claims.

These actors are often described as ‘powerful outlets’, lending a degree of authenticity that is otherwise difficult to replicate, but this review found little concrete assessment of impact for such programming. While the content itself may be erudite, it is difficult to determine effectiveness without an accurate understanding of who the messages are reaching and how they are being interpreted. Some studies simply provide descriptive overviews of ‘faith-based interventions’ with virtually no analysis of qualitative or quantitative results. In other instances, programmes have supplied ‘top-line findings’ without the methodological details and supplementary data for independent verification. Al Sakinah, an ‘edification and interactive communication initiative’ in Saudi Arabia, announced remarkably high success rates in 2008 for its efforts to contest extremist

---

40. P/CVE has subsequently extended its focus on other iterations of violent extremism – such as the far right – and the relevant studies have been examined in this review. However, most of the documentation primarily focused on ‘Islamist’ militancy.
42. Helmus, York and Chalk, ‘Promoting Online Voices for Countering Violent Extremism’.
43. Speckhard, ‘Battling the “University of Jihad”’.
45. Ibid.
46. Speckhard, ‘Battling the “University of Jihad”’.
ideology and debate sympathisers. Initially describing 877 individuals (722 male and 155 female) who were persuaded to renounce their ‘radical’ beliefs, more recent figures suggest the programme has ‘moderated the opinions’ and ‘corrected the views’ of 1,500 of the total 3,250 ‘radicals’ engaged in discussion. Forty percent of this cohort ‘have retreated completely from all or most of the radical thoughts they used to adopt ... the rest have abandoned the most dangerous of these ideas’. While studies burnish anecdotal examples of Al Sakinah’s impact by citing excerpts of transcripts between practitioners and recipient individuals from the campaign’s early outreach, it remains difficult to validate these claims or their underlying metrics.

The few comprehensive programme evaluations that are available, such as Lanny Octavia and Esti Wahyuni’s assessment of interventions to promote religious freedom and challenge extremist discourse in Indonesia, reveal uncertainty over whether these messaging activities actually achieved their specified P/CVE objectives. Broadcasting radio lectures by religious authorities and prominent ‘progressive Muslim figures’ across pesantrens (Islamic boarding schools) and satellite villages, accompanied by a series of youth-led media activities, the project was considered relatively effective based on feedback from key informant interviews, focus groups, and student and community surveys. However, the findings concede that the low frequency and limited transmission of these broadcasts constrained engagement with vulnerable individuals or conflict-afflicted localities. Instead, outputs were largely consumed by an already tolerant and peaceful audience, recipients predisposed to support such content. This is not to detract from the benefits of these initiatives, but it is unclear what impact they had on those susceptible to – or already enmeshed in – radicalisation or recruitment processes.

More broadly, studies question the effectiveness of directly refuting ideological stances and stress the dangers of accentuating ‘belief perseverance’. Joel Day and Scott Kleinmann, for

52. Ibid.
53. Ibid.
56. Ibid.
57. Ibid.
example, cite evidence of counterproductive ‘ideology-based CVE’ programmes that drive defensive audiences towards more radical positions as they increasingly feel under attack. Consequently, rebuttals of rigidly held positions could be rejected by the very audiences such campaigns should ostensibly be targeting, especially if the correction of propaganda or the advocacy of a ‘true reality’ is perceived to be condescending or dismissive by its recipients. Drawing on psychological and neuro-scientific explanations, Christina Nemr also queries the effectiveness of counternarrative and counter-ideology activities, citing the tendency of individuals to ‘push threatening information away in favour of information that confirms their own beliefs’. Treating P/CVE as a simple ‘belief-counter-belief’ dynamic and neglecting the emotional resonance that VE messaging sometimes elicits can therefore be dangerous, and runs the risk of unconsciously reinforcing the very attitudes it is trying to stifle.

Such approaches may similarly ignore the variation in recruitment processes, and the interpersonal connections and relationships that often precede the introduction of any political, ideological or religious proposition. Manni Crone, for instance, distinguishes cognitive from behavioural extremism and disputes any proposed linearity between the two, suggesting a commitment to radical ideology is not necessarily a precondition for violence. Conflict, criminality or more banal, non-ideological motivations can often pave the way for terrorist

60. Van Eerten et al., ‘Developing a Social Media Response to Radicalization’.
63. Ibid.
engagement. These could therefore be considered ‘equifinal’ dynamics, with an eclectic set of pathways leading to the same end state. Beliefs are of course relevant, but may play several roles requiring discrete responses, operate ‘differently for different people’ and create highly variegated processes unrelated to any immediate shift in ideological outlook. Likewise, in a comprehensive review of counter-messaging documentation in 2016, Kate Ferguson found little evidence that exposure to malicious material induces violent activity. The broader literature on hate speech identifies a possible link between ‘violent words and violent deeds’ in sectarian, ethno-nationalist and xenophobic contexts, but this evidence is still in a ‘relatively raw form’ and only indicates mass-level trends. While a narrow focus on ideological factors might appear to be a convenient avenue for intervention, it risks overlooking the factors that actually foster appetite for extremist material in the first place, and neglects a web of personal motivations, structural drivers and the social bonds (online and offline) forged through recruitment. Recent additions to the radicalisation literature – such as Noémie Bouhana’s ‘SSS inference framework’ – reinforce this point, casting exposure to an extremist milieu and content as a single dimension of a much more complex process.

Even on a discursive level, contesting ideological claims appears insufficient when conducted in isolation, in part because interventions often misconstrue or inadequately interrogate how target audiences interpret information. Anne Aly specifies a lack of academic studies examining how audiences ‘decode’, and sometimes act on, the messages embedded in VE propaganda. Groups

74. Ferguson, ‘Countering Violent Extremism Through Media and Communication Strategies’.
75. Ibid.
76. Ibid.
79. Davies et al., ‘Toward a Framework Understanding of Online Programs for Countering Violent Extremism’.
like Al-Qa’ida have developed a ‘single’ or ‘meta-narrative’ – a framework that integrates political, moral, religious and socio-psychological elements. As a result, VE messages – conditioned to specific contexts – can leverage different thematic combinations that tap into the heterogeneous interests of a fairly broad audience, at least on some level. Under such conditions, it becomes difficult to ‘reverse engineer’ violent strands of communication as they intersect with, and feed into, an ‘amorphous and tangled fabric’ of identities, relationships, interests and understandings. As Christian Leuprecht and colleagues likewise explain, ‘different parts of the Global Jihadist narrative are held by different audiences and each part, its prospective audience and the enabling mechanisms of radicalisation must be targeted separately for counter-narratives to be effective’.

In their theoretical prescriptions for constructing and disseminating counternarratives, Kurt Braddock and John Horgan acknowledge the importance of responding to specific ‘incongruities and contradictions’ embedded in extremist content, which may involve theological arguments and contextualising scriptural references, but reference calls for a more comprehensive system of stories. Unfortunately, many P/CVE interventions tend to disrupt or refute particular segments and threads of VE narrative but struggle to ‘target every stage of engagement’, particularly as they often rely on standalone messages.

Additional challenges apply to the promotion of ‘moderation’ – a subjective and contentious term, particularly when sanctioned by government, that has precipitated some resistance as stakeholders, including pre-eminent grassroot leaders, protest against the perceived ‘criminalisation of ideology’. This has led to increasing reticence of religious and secular

83. Sedgwick, ‘Jihadist Ideology, Western Counter-Ideology, and the ABC Model’.
84. Davies et al., ‘Toward a Framework Understanding of Online Programs for Countering Violent Extremism’.
institutions, NGOs and civic activists to ‘wade into matters of ideology’ and risk becoming branded as ‘sell-outs’.

Similarly, the credibility and access of moderate and mainstream voices has been questioned, with programme evaluations in Kenya indicating that clerical authorities occasionally lacked the knowledge and experience to speak fluently on topics of interest to younger constituents such as current affairs and foreign policy issues. Research suggests these stakeholders may not only be rather marginal across the social networks of vulnerable populations in certain areas, but potentially even feature in the factors contributing towards radicalisation and recruitment. One programme evaluation highlighted the tendency of older religious and community leaders to monopolise positions on decision-making structures like local mosque committees in various localities, relegating youth voices and incentivising the exploration of alternative, potentially more radical options. These dynamics are context-dependent and by no means universal, but they do call into question the assumed agency and authority of ‘moderate actors’ in relation to P/CVE activities. The inclusion of comparatively extreme but non-violent figures may be more palatable to ‘at-risk’ audiences, or at least those individuals further downstream in the radicalisation and/or recruitment process, but they are usually controversial, requiring the navigation of thorny ethical, political and strategic questions. Governments, for instance, appear reticent to collaborate with such stakeholders for fear of being seen to legitimise forms of ‘extremism’.

Moreover, the capacity of ‘moderate ideas’ to either out-compete or gradually assimilate more fringe ideologies has only limited support in the literature. For example, Rogelio Alonso’s historical analysis of separatist militancy in Spain suggests nationalists can potentially damage terrorist groups such as ETA (Euskadi Ta Askatasuna) at their ‘root and heart’ by normalising non-violent strains of political activity. The study partially attributes the increasing unpopularity of ETA’s terrorism to a counter-messaging campaign that helped bolster more traditional movements.

90. Ibid.
95. RUSI, ‘STRIVE’.
such as the Basque Nationalist Party (PNV). However, there is an absence of explicit causal chains between this strategy and ETA’s gradual disengagement, making it difficult to substantiate conclusions.\textsuperscript{99} Other studies suggesting mainstream political conservatives may have influence over more radical white nationalists draw on untested hypotheses and theoretical inferences.\textsuperscript{100}

In contrast, digital interventions such as ‘Abdullah X’ and ‘Average Mohammed’ could potentially be more promising as they critically explore the interests, anxieties and frustrations of ‘vulnerable individuals’ rather than engaging in a direct rebuttal. This aligns with a widespread claim in the P/CVE field that Muslims in diasporas are ‘arguably the best social carriers to counter al Qaeda’s single narrative’, at least in a Western context, as they can ‘create a new meaningful and credible’ message expressing their ‘vision’ of what it means to be a contemporary believer.\textsuperscript{101} Delivering content through a series of cartoon videos on YouTube, ‘Abdullah X’ discusses ‘legitimate grievances’ and frank questions surrounding Muslim identity in the UK, the Syrian Civil War, normative Muslim values and the Islamic State, while simultaneously ‘admonishing’ the use of violence and calling for a ‘redirection’ of this energy.\textsuperscript{102} ‘Average Mohammed’ similarly produces videos that address ‘identity, gender equality, democracy, being Muslim in a Western culture and slavery’ through the medium of animation.\textsuperscript{103} In this sense, they ‘speak from a religious perspective which appeals to the same motivational, normative and aesthetic values as Islamic State propaganda but re-embeds these in the modern world’.\textsuperscript{104} Such approaches help dissolve the binary comparisons that often underpin extremist narratives,\textsuperscript{105} severing the links groups like the Islamic State draw between themselves and their recruits against various ‘out-groups’.\textsuperscript{106} As James Digby summarises, these interventions therefore ‘retain the powerful moral, epistemological and aesthetic vision provided by Islamic State’s Salafi-Jihadist narratives while undermining the conclusions they reach about normative action’.\textsuperscript{107}

Despite this promise, there are limits to how far any ‘success’ is corroborated by the empirical evidence. Both projects cultivated an average view count of a ‘few thousand per video’—numbers

\begin{small}
\textsuperscript{99} Ibid.
\textsuperscript{100} Berger and Strathearn, ‘Who Matters Online’.
\textsuperscript{101} Schmid, ‘Al-Qaeda’s “Single Narrative” and Attempts to Develop Counter-Narratives’, p. 28.
\textsuperscript{102} Jonathan Russell, ‘Developing Counter- and Alternative Narratives Together with Local Communities’, RAN Issue Paper, Radicalisation Awareness Network (RAN), 15 October 2018; Beutel et al., ‘Guiding Principles for Countering and Displacing Extremist Narratives’.
\textsuperscript{104} Digby, ‘Cartoons vs. the Caliphate’.
\textsuperscript{105} Braddock and Horgan, ‘Towards a Guide for Constructing and Disseminating Counternarratives to Reduce Support for Terrorism’.
\textsuperscript{106} J M Berger, ‘Deconstruction of Identity Concepts in Islamic State Propaganda: A Linkage-Based Approach to Counter-Terrorism Strategic Communications’, Europol, 9 June 2017, cited in Digby, ‘Cartoons vs. the Caliphate’.
\textsuperscript{107} Digby, ‘Cartoons vs. the Caliphate’, p. 71.
\end{small}
described as neither ‘insignificant nor staggering’, described as neither ‘insignificant nor staggering’, and not necessarily commensurate with the attention they receive from Western governments. This does not negate the value of these activities, but it is difficult to assess how much of an impact they have actually achieved in countering radicalisation and recruitment, especially as no new content has been produced for ‘Abdullah X’ since October 2016 and many of these campaigns now appear ‘largely defunct’. Nonetheless, the underlying logic of their approach chimes with the practical experience of Toaha Qureshi and Sarah Marsden, who describe the value of discussion of day-to-day questions, problems and anxieties, investment in an ‘ongoing dialogue’, and the promotion of an alternative set of attitudes ‘founded in authentic Islamic scholarship, social integration and personal responsibility’. As described later in this paper, integrating conversations that draw on the same moral and epistemological foundations as VE groups into a wider set of educational, social and supportive activities could potentially have traction as it helps substantiate a more comprehensive and dynamic set of messages with lived experience.

**Scale, Resonance and ‘Message Dominance’**

It is clear from the literature that some interventions fail to resonate due to problems with the message itself. The most frequently cited example is the US State Department’s Centre for Strategic Counterterrorism Communications (CSCC), which co-opted Islamic State imagery and exposed the violent realities of life in the group’s territory in a campaign that was widely criticised as counterproductively amplifying ‘the most effective dimensions of [the Islamic State’s] marketing campaign’. A 2015 external review dismissed these outputs as ineffective, particularly when competing with extremist propaganda, and academic studies criticised their

---

108. Ibid., p. 3.
112. Ibid.
'leaden' and 'amateurish' appearance.\textsuperscript{115} They not only tend to have low production values but reflect inappropriate aesthetic and musical conventions that may alienate the target audience, in contrast to the powerful use of anashid (a cappella songs) by groups like the Islamic State to enhance their discursive content and video footage.\textsuperscript{116} This leaves US officials at a disadvantage as they cannot leverage these techniques without seeming either clumsy or inappropriate.\textsuperscript{117} While papers suggest that music promoting ‘tolerance and reconciliation could be used to reconnect with at-risk individuals’ and compete with the emotional investment often elicited by VE content, little research or empirical data is currently available to substantiate these claims or the agency of ‘local and regional music scenes and sub-cultures’.\textsuperscript{118}

Other studies criticise a prevalent strategy of mass persuasion in counter-messaging, particularly in the online space, describing a reliance on ‘message dominance’ that involves ‘rebutting or purging antithetical viewpoints while allocating the majority of resources to actively spreading the message’.\textsuperscript{119} This focus is not surprising given the scope of the internet and Web 2.0 social media platforms as both an operational domain and a ‘marketplace of ideas’.\textsuperscript{120} Dynamic ‘prosumption’ processes\textsuperscript{121} have accelerated and decentralised digital output, generating new forms of activism\textsuperscript{122} and galvanising competition between state, alternative and insurgent ‘truth perspectives’.\textsuperscript{123} This has helped democratise information creation in ways that contest conventional religious and political brokerage,\textsuperscript{124} and streamline access to radical ‘ideological structures’ that provide ‘some semblance of order, a renewed sense of self, and greater meaning to troubled realities’.\textsuperscript{125} Under these conditions, the motivation to contest VE propaganda is


\textsuperscript{116} Bean and Edgar, ‘A Genosonic Analysis of ISIL and US Counter-Extremism Video Messages’.

\textsuperscript{117} \textit{Ibid.}


\textsuperscript{119} Freear and Glazzard, ‘Preventive Communication’, p. 92.

\textsuperscript{120} Neumann, ‘Options and Strategies for Countering Online Radicalization in the United States’.

\textsuperscript{121} Where stakeholders are simultaneously consumers and producers of content.


\textsuperscript{123} H L Goodall Jr et al., ‘Rhetorical Charms: The Promise and Pitfalls of Humor and Ridicule as Strategies to Counter Extremist Narratives’, \textit{Perspectives on Terrorism} (Vol. 6, No. 1, March 2012), p. 74.

\textsuperscript{124} Aly et al., ‘Making “Noise” Online’.

understandable and contributes to the urgency of ‘message dominance’ due to the density and accessibility of malicious content online. The response has assumed various forms, with interventions – particularly upstream preventive outreach – expanding the circulation of ‘positive content’ and counternarratives through the use of algorithms on social media;\(^\text{126}\) others launching ‘social marketing’ and ‘Trojan style’ interceptions to amplify evidence-based counter-arguments;\(^\text{127}\) and a number of projects trying to experiment with different methods of online influencing.

However, these approaches often struggle in three important areas: targeting the right audience; translating reach into resonance; and engaging the wider social dynamics that frequently drive VE recruitment and inform how consumers interpret and respond to content.

First, on a practical level, the anonymity of many social media account holders makes it difficult to verify whether the audiences consuming P/CVE content are actually those that should be targeted.\(^\text{128}\) Mass-marketing strategies can lead to generic content with little resonance among those genuinely ‘at risk’ or a dependence on ineffective ‘scatter-gun approaches’.\(^\text{129}\) Conversely, automation may help identify online profiles expressing support for extremist ‘talking-points’,\(^\text{130}\) but focusing on ‘narrative radicals’ could be misguided as studies suggest those more likely to act violently do not usually engage in conversation.\(^\text{131}\) This can lead to the prospect of ‘false positives’ that prioritise free riders rather than ‘action radicals’.\(^\text{132}\) As Maura Conway summarises:

> how can we develop and deploy effective online CVE projects absent knowing who precisely these should be targeted at, what types of content are attractive to them, and what platforms are trafficked by these users? These are the kinds of answers that nobody appears to have at the present time.\(^\text{133}\)


\(^{129}\) Hemmingsen and Castro, The Trouble with Counter-Narratives; Shandon Harris-Hogan, Kate Barrelle and Andrew Zammit, ‘What is Countering Violent Extremism? Exploring CVE Policy and Practice in Australia’, Behavioural Sciences of Terrorism and Political Aggression (Vol. 8, No. 1, 2016), pp. 6–24.

\(^{130}\) Davey, Tuck and Amarasingam, ‘An Imprecise Science’.

\(^{131}\) Leuprecht et al., ‘Containing the Narrative’.

\(^{132}\) Ibid.

\(^{133}\) Conway, ‘Determining the Role of the Internet in Violent Extremism and Terrorism’, p. 82.
Consequently, it often becomes difficult to tailor and contextualise interventions\textsuperscript{134} – shortcomings that can have significant implications, as described by Stevie Voogt in their account of ‘Community Action for Preventing Extremism’ (CAPE), an online counter-messaging project tackling far-right recruitment in Australia. Despite an overlapping membership, traditional white supremacists and local anti-Muslim organisations have qualitatively distinct ideas, discourses and motivations.\textsuperscript{135} Without accurately delineating between the two, counternarratives directed towards one group may not be relevant to, or could even strengthen, the arguments of the other.\textsuperscript{136} These implications cut across various forms of counter-messaging, with theoretical analysis of satirical content suggesting that humour can become a ‘double-edged sword’ if it is not adequately refined, targeted or sensitised to the cultural norms of the respective audience.\textsuperscript{137} The direct ridicule of VE ideas can be counterproductive,\textsuperscript{138} eliciting ‘back-fire effects’ that increase consumer defiance and engrain in-group mentalities.\textsuperscript{139} Without understanding the disposition of those receiving, engaging and responding to such outputs, practitioners therefore risk ‘preaching to the choir’ and missing vulnerable populations entirely, or worse, generating harmful outcomes. The use of netnography to understand social interaction in digital spaces may alleviate some of the wider issues of targeting and audience identification by helping to map specific online subcultures and hallmark content, but this research remains in its infancy.\textsuperscript{140}

Second, in many cases it is unclear whether ‘reach’ equates to resonance. For example, pilot schemes like Extreme Dialogue experimented with commercial and organic methods for distributing counternarratives as advertisements across different countries.\textsuperscript{141} Targeting ‘at-risk individuals’ identified through their browser cookie data, the project circulated short, emotive films of former extremists, victims of terrorism, frontline workers and other stakeholders recounting their personal experiences on platforms like YouTube, Facebook and Twitter\textsuperscript{142} to help dispel ‘extremist myths and encourage empathy with and understanding of the Other’.\textsuperscript{143}

\textsuperscript{134} Studies describing a clear lack of contextualisation in counter-messaging include Davies et al., ‘Toward a Framework Understanding of Online Programs for Countering Violent Extremism’; Naomi Theuri, ‘Gender and Contextual Perspective in Countering Violent Extremism (CVE): Examining Inclusion of Women and Contextual Factors in Online Approaches to CVE’, degree project, Department of Criminology, Malmö University, 2017.


\textsuperscript{136} Ibid.

\textsuperscript{137} Beutel et al., ‘Guiding Principles for Countering and Displacing Extremist Narratives’.

\textsuperscript{138} Anonymous, ‘Tackling White Supremacy, Australian Style’.

\textsuperscript{139} Day and Kleinmann, ‘Combating the Cult of ISIS’; Lee, ‘Informal Countermessaging’; Beutel et al., ‘Guiding Principles for Countering and Displacing Extremist Narratives’.

\textsuperscript{140} Netnography is an online research method drawing on ethnography that helps map and understand social interactions in digital communication contexts; Waldman and Verga, ‘Countering Violent Extremism on Social Media’; Greenberg, ‘Counter-Radicalization Via the Internet’.

\textsuperscript{141} Waldman and Verga, ‘Countering Violent Extremism on Social Media’.

\textsuperscript{142} Reynolds and Tuck, ‘The Counter-Narrative Monitoring and Evaluation Handbook’.

\textsuperscript{143} Helmus and Klein, ‘Assessing Outcomes of Online Campaigns Countering Violent Extremism’.
This led to quantitatively impressive results, with the Canadian strand alone receiving 504,000 ‘impressions’ on Facebook, and accumulating 50,673 views on YouTube in its first week.\textsuperscript{144} Despite reaching a wider audience of high-risk individuals, paid advertising nevertheless failed to generate sustained interest as any effects quickly dissipated after the first week of the campaign,\textsuperscript{145} underscoring the difference between the level of exposure to, and the effectiveness and resonance of, counter-messaging.\textsuperscript{146}

Other interventions have struggled to verify whether their online outreach has precipitated behavioural change or any shift in consumer attitudes offline.\textsuperscript{147} The International Center for the Study of Violent Extremism developed 19 advertisement campaigns to engage ‘at-risk’ Iraqi nationals over Facebook, using two films to expose the hard realities of life in the ‘caliphate’ and convert frustration into a ‘positive exchange of ideas and participation’.\textsuperscript{148} Labelled with pro-Islamic State or ambivalent names, these videos were designed to permeate and infect the ‘digital territory’ of VE groups. Speckhard and colleagues highlight a cumulative reach of over 1.2 million generated over 24 days, achieving approximately 1.7 million views and 2.3 million impressions, but concede their metrics have limitations due to the experimental nature of the project.\textsuperscript{149} Online comments were cited as an indicative benchmark for ‘influence’, although this is a self-selective exercise that typically involves a small minority of users, and may not accurately reflect the positions or proclivities of their authors.\textsuperscript{150} As a result, this review found these findings to be largely inconclusive in relation to the campaign’s impact on those receiving the messages.

While programmes have tried to compensate for these methodological shortfalls by emphasising the inherent value of ‘sowing the seeds of doubt’, the near-ubiquitous lack of substantiating

\textsuperscript{146}. It should be noted that this social media outreach was not the sole focus of Extreme Dialogue. It also provides resource packs and educational content for teachers, although this review found little publicly available evaluation material assessing its wider effectiveness and impact.
\textsuperscript{148}. Anne Speckhard et al., ‘Mounting a Facebook Brand Awareness and Safety Ad Campaign to Break the ISIS Brand in Iraq’, Perspectives on Terrorism (Vol. 12, No. 3, June 2018), pp. 50–66. Similar pilot projects were also conducted with other online demographics such as target populations in the Western Balkans. See Anne Speckhard, Lorando Bodo and Haris Fazliu, ‘Bringing Down the Digital Caliphate: A Breaking the ISIS Brand Counter-Narratives Intervention with Albanian Speaking Facebook Accounts’, Research Report, International Center for the Study of Violent Extremism, 2018.
\textsuperscript{149}. Ibid.
\textsuperscript{150}. Helmus and Klein, ‘Assessing Outcomes of Online Campaigns Countering Violent Extremism’. 
data hampers an understanding of whether ‘proto-typical influencing efforts actually work’ and often leave practitioners simply relying on ‘gut instinct’. Few reports enumerating the success or failure of such interventions ‘go beyond the number of impressions recorded by any given message … which do not necessarily correlate to the prevention of VE activity’. In a review of prominent counternarrative campaigns, Todd C Helmus and Kurt Klein found none with any assessment of impact. Even comprehensive analysis of pilot counter-messaging activities like Average Mohammed or Harakat-ut-Taleem (HUT) – a local organisation countering Taliban recruitment narratives in Pakistan – largely relied on output-level data and indices such as reach, retention rates and sustained engagement, which struggle to gauge or proxy higher-level impacts or mitigate the distortive activity of online bots, as acknowledged by the practitioners themselves. These appraisals may therefore be useful for improving the technical capabilities and marketing strategies of interventions, but it often remains unclear whether such approaches actually resonate and persuade those susceptible to radicalisation and recruitment.

Despite these issues, the importance of strengthening organisational capabilities to help improve the quality, reach and scale of counter-messaging efforts is heavily emphasised in the literature given the demand for greater plurality and heterogeneity in counternarrative material. Many authoritative messengers lack the ‘networks, support and specialist expertise’ needed to increase their activities and connect with their intended audience. In this context, various training schemes have proven relatively effective. For instance, a detailed evaluation of Youth Innovation Labs – a youth-focused capacity-building initiative – suggested it achieved its primary objectives of increasing the reach and production value of third-party content. A media training programme in Morocco led by Search for Common Ground enumerates similar accomplishments, helping to amplify locally produced alternative narratives that far surpassed their anticipated benchmarks for consumer interaction. In Kenya, 54 civil society organisations (CSOs) were supported by the Institute for Strategic Dialogue (ISD) in the creation

151. Ibid.
155. Defined as individual users writing more than one comment or sending a direct message to the host organisation.
of counternarratives, with 29 participants developing their own campaigns.\textsuperscript{161} Manufacturing videos to address both VE and political violence, the interventions were collectively able to reach 4.4 million users – over 10\% of Kenyans online – during the 2017 presidential election.\textsuperscript{162} This expanded circulation suggests proactive specialist-led training can drastically improve the volume and scope of counter-messaging.\textsuperscript{163}

Such programmes clearly benefited participants and produced positive effects, from improving the substance of messaging activities to boosting their dissemination. However, assessment metrics for tracking the intended results of the broader counter-messaging campaigns themselves still relied on output-level data primarily oriented around reach and viewership, providing little data on whether the ‘amplification of local voices’ actually had any ameliorating impact on the wider problem of VE. For example, ‘The Truth About Halal’ – a CSO-led film nurtured by Youth Innovation Labs – was described as ‘well received’, collating over 38,000 views in two weeks.\textsuperscript{164} Yet, this information gives little insight as to the audience’s demographic composition, making it difficult to determine whether the content resonated with those susceptible to radicalisation and recruitment or had any effect on their outlook and behaviour.

Third, the overarching preoccupation with ‘message dominance’ seems in some cases to resemble aspects of the ‘hypodermic’ or ‘magic bullet’ model of communication, which assumes that people respond directly to communications content. Cristina Archetti has described this paradigm as widespread in P/CVE communication-based approaches, which tend to cast radicalisation as a passive process of indoctrination and imply that messages can therefore be ‘injected into society to inoculate individuals against external influence’.\textsuperscript{165} However, this premise has been challenged both empirically and theoretically. There is little evidence of a linear relationship between exposure to online extremist messaging and violent behaviour, and notions of ‘self-radicalisation’ have been broadly disputed.\textsuperscript{166} Instead, the role of offline and virtual social networks appear to be paramount in encouraging VE conformity, groupthink and

\begin{flushleft}
\textsuperscript{162} Ibid.
\textsuperscript{163} Ibid.
\textsuperscript{164} Saltman, Dow and Bjornsgaard, ‘Youth Innovation Labs’.
\end{flushleft}
mobilisation, with exposure to extreme content often ‘reinforcing already made up minds’.\textsuperscript{167} This is supported by communications theory, which indicates that ‘narrative is not a message to be sent out to an audience in order to trigger certain expected (and predictable behaviours)’, but a social product informed by ‘socialisation, meaning-seeking and identity formation’.\textsuperscript{168}

The salience of VE material derives, at least in part, from the networks that convey it, with Day and Kleinmann arguing that organisations like the Islamic State exploit affective bonds and friendship ties as channels for circulating their doctrine and imbuing it with meaning.\textsuperscript{169} Trying to ‘inoculate’ target audiences by injecting counter-arguments of any flavour without the wider social infrastructure to infuse them with relevance and resonance may therefore constrain the influence of these messages, or may even ‘end up confirming radical views’.\textsuperscript{170}

Nevertheless, of social media projects reviewed by Garth Davies and colleagues, for example, none had a significant interest in the social processes and broader factors driving radicalisation.\textsuperscript{171} The study also notes a broader lack of socio-psychological considerations in many online interventions and the limitations of ‘refuting erroneous ideological interpretations, correcting historical inaccuracies and unravelling conspiracy theories’ without ‘offsetting or replacing’ the relational dynamics and support networks accompanying VE outreach.\textsuperscript{172} In only focusing on the message, counternarratives not only miss crucial mechanisms that often make VE content persuasive, but overlook the personal connections and social linkages actually facilitating and expediting VE recruitment – dynamics typically transcending discursive content.\textsuperscript{173} Interactive engagement, dialogue and networked approaches have been tested as possible solutions – and are analysed below as promising approaches – but many of the 60 collected studies focusing on online interventions seem to omit these important faculties. Until such gaps are resolved, both the resonance and scope of counternarratives will likely remain insufficient.

\textsuperscript{169} Day and Kleinmann, ‘Combating the Cult of ISIS’.
\textsuperscript{171} Davies et al., ‘Toward a Framework Understanding of Online Programs for Countering Violent Extremism’; Van Eerten et al., ‘Developing a Social Media Response to Radicalization’, p. 49.
\textsuperscript{172} Davies et al., ‘Toward a Framework Understanding of Online Programs for Countering Violent Extremism’, p. 79.
Credible Messengers and the Role of the State

Governments and public institutions are often considered compromised sources of counter-messaging in the eyes of target audiences, with state-led campaigns receiving criticism for appearing clumsy, ineffective or trying to ‘stifle debate’. CSCC’s ‘Think Again Turn Away’ campaign is, again, a widely cited example, not least as it inadvertently endowed Islamic State sympathisers with credibility and a platform. Members of CSCC expressed discomfort at sending counter-messages with State Department branding, and described how it became a locus for public ridicule that seemingly lacked ‘even the most basic understanding of the complex conflict’ or the affordances of digital communication.

This lack of messenger credibility is evident even for programmes that otherwise display a sophisticated approach to the content of messaging campaigns – such as ‘Say No to Terror’ (SNTT), a counter-messaging campaign predominantly targeting audiences in Saudi Arabia. Drawing on Islamic symbolism, scripture and the cultural norms of the target audience, SNTT packaged themes that emphasise the criminality of terrorist groups, the hypocrisy of their narratives and the validity of ‘traditional’ religious models and authority structures. Crucially, it appealed to both the rational and emotional dimensions of recruitment, developing messages that leveraged heuristic and systematic persuasion. Consequently, evaluators have applauded the mechanics and methodology of this approach, but criticised its execution. Using 281 YouTube comments as a proxy for public perception, Aly and colleagues identify 60% as ‘negative’. Much of this feedback focuses on SNTT’s origins and lack of transparency. Its sponsorship and stakeholders are unknown, leading to allegations of ‘astro-turfing’ – a tactic of artificially mimicking a grassroots

---

176. Davies et al., ‘Toward a Framework Understanding of Online Programs for Countering Violent Extremism’.
178. Reed, Ingram and Whittaker, ‘Countering Terrorist Narratives’.
180. Van Eerten et al., ‘Developing a Social Media Response to Radicalization’.
181. ‘Noise’ is defined as a method for disrupting messages and preventing them from achieving their desired effect. See Aly et al., ‘Making “Noise” Online’.
183. Ibid.
As such, the project is considered by some, including many of its recipients, to be affiliated with various state sponsors, diminishing trust in its activities and compromising the delivery of its message, irrespective of its quality or relevance.\textsuperscript{185}

Studies concede the value of states supporting public discourse around ‘anti-exclusion and to counter the manipulation of religion to justify violence’, but this is often overshadowed by their ‘securitisation of [P/CVE] campaigns’.\textsuperscript{186} Instead, government ‘calls to action’ may be ‘more appropriate than trying to shape opinions’,\textsuperscript{187} with analysis emphasising the importance of using states’ financial resources, technical specialisms and infrastructure to help enhance the agency of independent stakeholders and critical ‘third party voices’ such as local CSOs and community stakeholders.\textsuperscript{188} Projects discussed in the literature include a series of web-based seminars delivered by the US State Department to facilitate: networking and co-creation between ‘moderate’ Somali bloggers in Europe, Canada and Africa;\textsuperscript{189} training in online marketing strategies; and initiatives such as ‘Viral Peace’, which adopted a similar template of enabling grassroot stakeholders to seed and maximise outreach by strengthening the skills of ‘key influencers’ in social media, narrative creation and message dissemination across South Asia.\textsuperscript{190} Accounts of these programmes are generally positive given the emphasis placed on augmenting organic, locally generated content. However, many assessments provide cursory descriptions rather than any analysis into what the sessions achieve in relation to P/CVE objectives. The State Department’s social media training scheme was ‘judged a success’ on the basis of participants’ enthusiastic reception to the course – a procedural accomplishment that otherwise offers little insight into the programme’s results or longer-term outcomes.\textsuperscript{191} Other initiatives sponsored

\begin{footnotesize}
\textsuperscript{184} Ibid., p. 43; Van Eerten et al., ‘Developing a Social Media Response to Radicalization’, p. 29.
\textsuperscript{188} Abbas Barzegar, Shawn Powers and Nagham El Karhili, ‘Civic Approaches to Confronting Violent Extremism: Sector Recommendations and Best Practices’, Georgia State University, British Council and ISD, September 2016; Russell, ‘Developing Counter- and Alternative Narratives Together with Local Communities’.
\textsuperscript{189} Neumann, ‘Options and Strategies for Countering Online Radicalization in the United States’, p. 446.
\textsuperscript{191} Helmus, York and Chalk, ‘Promoting Online Voices for Countering Violent Extremism’, p. 6.
\end{footnotesize}
by the ‘Preventing Violent Extremism Pathfinder Fund’ in the UK have little evaluation material publicly available.  

Some capacity-building approaches also tend to prioritise technical refinements to the content of messages over investing in the infrastructure and social capital of religious and civic institutions. Studies cite the need to encourage resilience and deeper community ties, developing the capacity of influential leaders and local stakeholders, expanding their networks and marketing a holistic alternative to violence through ‘affective bonds, social practices and friendships’. Yet, these components can frequently be overlooked in conventional training schemes, with many failing to explicitly connect discursive activities with other forms of outreach.

Crucially, close cooperation between government and CSOs may likewise diminish the credibility and legitimacy of the latter: even in countries with comparably mature civil traditions, numerous independent stakeholders have struggled with perceptions of being front organisations.

Beyond these programmatic limitations and risks, there are wider questions over the role of conventional CSOs as messengers that are unresolved in the literature. Studies generally indicate the importance of social networks, describing their capacity to transmit ideas across their membership. Even in underdeveloped rural environments or communities with weak societal bonds, these linkages can quickly facilitate the spread of new behaviours and attitudes through learning and peer-to-peer influencing. However, network analysis reveals civil society activists are fairly peripheral in some of these social circuitries. In assuming CSOs function homogenously, enjoy access to vulnerable populations and should generally lead interventions, there is a risk that reductive parameters of ‘legitimate’ P/CVE activity will be imposed which do not always cohere with local realities. Informal authorities, friends, associations and acquaintances are key points of contact for engaging vulnerable groups, but these stakeholders are not necessarily included under the CSO umbrella or engaged in traditional outreach efforts.

Moustafa Ayad and Amarnath Amarasingam, alongside others, specifically argue for the need to integrate young stakeholders across ‘every stage of the design, development and promotion

193. Day and Kleinmann, ‘Combating the Cult of ISIS’.
196. Ibid.
197. Russell, ‘Meet Me at the Maskani’.
of campaigns’ to ensure more effective, resonant and authentic programming.199 Their absence is evident in many of the interventions evidenced in the literature, such as Denmark’s Dialogue Forums (PET). These sessions provide discussion opportunities, reciprocal learning and lateral networking between state, security and civilian stakeholders that have ‘insight and influence’ in their neighbourhoods, in part to better enhance communication and outreach efforts.200 While ‘the actual impact of these countermessages on prevention is unclear and needs further study’, Julian Brett argues this approach is likely to build the capacity of beneficiaries to identify ‘at-risk’ people, and improve the quality and granularity of messaging activities.201 However, the primary recipients of such content – young people – are not represented in PET,202 raising questions over the benefits of collaborative schemes that prioritise ‘traditional’ P/CVE actors but exclude those individuals that may have credibility, trustworthiness and access to persuade others in their peer networks.

Certain donors have tried to alleviate such issues through competitions and grant schemes, encouraging greater grassroots innovation and autonomous youth-led creativity. ‘Peer 2 Peer’, for example, is repeatedly cited as a ‘success story in government cooperation with the private sector’;203 enabling students to design and disseminate ‘anti-hate’ and P/CVE messaging.204 This outreach undoubtedly expanded the capacity of local stakeholders across different contexts,205 but there are few evaluations that rigorously assessed impact on wider recruitment dynamics.206 More broadly, little research has been conducted into the capacity and conditions for effective youth agency in the prevention space. ‘Youth’ as a working concept is rarely defined and studies tend to neglect the intersectionality and distinctions between ‘young people’ as a generalised

201. Ibid., p. 5.
202. Ibid., p. 4. The evaluation does suggest this limitation may be caveated by the inclusion of individuals with ‘good levels of access’ to youth groups, although the study is unclear whether this is the case.
204. Szmania and Fincher, ‘Countering Violent Extremism Online and Offline’; Digby, ‘Cartoons vs. the Caliphate’.
category and those susceptible to recruitment. This can sometimes make it difficult to achieve the full potential of empowerment schemes and risks the creation of a ‘cottager industry’ with ‘more form than function’.

Independent websites, chat forums, viral media content and support groups are also relatively untapped resources in the prevention space, despite often displaying greater flexibility, relevance and popularity than their ‘formal’ P/CVE counterparts. Studies stress the need to harness and coordinate this ‘natural world content’ to help satiate demand for high-quality material without compromising its authenticity and creativity. Highlighting the close cooperation of practitioners and private citizens in the cyber security realm, Benjamin Lee – among others – references the value of leveraging the symbolic capital and legitimacy exercised by ‘informal content creators’. However, it is unclear what this means in terms of practical next steps, and there have been limits on how far INGOs and governments have successfully collaborated with these voices. For instance, ‘politically charged Hip-Hop’ has emerged as a relatively successful medium for alternative messaging across Muslim civil societies, offering an organic form of ‘transnational youth culture’ that enables ‘creative self-expression and cultural critique’. Unfortunately, governments are ‘cautious in supporting such projects for fear of losing control of the message’, leaving the P/CVE industry largely dependent on the ‘usual suspects’. The possibility of securitising, co-opting or tarnishing ‘natural content’ is likewise a constraint. It is also important not to exaggerate the authority and impact of this material. For example, empirical analysis suggests ‘counter-speech’ – a ‘common crowd-sourced response to extremism’ on Facebook – can display significant flexibility and capability. However, these approaches were ‘not always as effective as [they] could be’, given a large proportion of the reviewed messaging was deemed ‘non-constructive’ (abuse or aggressive threats). In contrast, satirical and comedic material was considered broadly popular but – as previously highlighted – risks alienating ‘vulnerable’ populations if they are not tailored appropriately.

212. Ibid.
214. Ibid.
215. Richardson, ‘Fighting Fire with Fire’.
216. Jamie Bartlett and Alex Krasodomski-Jones, ‘Counter-Speech Examining Content That Challenges Extremism Online’, Demos, October 2015.
Some pilot approaches – such as the ‘Redirect Method’, a commonly cited case study across the literature – have tried leveraging this material, but it remains unclear how effective they have been. Developed by a private consortium led by Jigsaw, a subsidiary of Alphabet Inc,\(^\text{218}\) the project integrated Google’s AdWord technology to peg counter-messaging advertisements to search terms associated with prominent VE narratives.\(^\text{219}\) Footage was curated from existing videos in English and Arabic featuring citizen journalism, documentaries, religious debates and testimonies from defectors\(^\text{220}\) to help ‘nudge’\(^\text{221}\) vulnerable individuals away from both jihadist and, later, far-right influences.\(^\text{222}\) Although an estimated 320,906 users actively ‘clicked’ on these clips in the first two months, and a technical guide for derivative campaigns was subsequently released, determining what impact it actually achieved in relation to radicalisation and recruitment remains difficult.\(^\text{223}\)


\(^{221}\) Reed, Ingram and Whittaker, ‘Countering Terrorist Narratives’.

\(^{222}\) Helmus and Klein, ‘Assessing Outcomes of Online Campaigns Countering Violent Extremism’.

II. What Looks More Promising?

This chapter identifies and analyses strands of P/CVE programming that have a stronger theoretical basis or which appear to have delivered more promising results. These activities generally: recognise and respond to the sociological and relational dynamics of radicalisation and recruitment; use credible voices to engage in interactive dialogue; and deliver integrative approaches that employ communications strategically, supplement messaging with wider activities and prioritise process.

Improving Critical Thinking and Media Literacy

Given the difficulties of identifying vulnerable individuals and creating palatable material to counter or disrupt VE propaganda, media literacy and critical thinking have been described as useful for enhancing self-regulation and individual resilience. Providing young people with the skills they need to safely navigate social media has already proven to be relatively effective in mitigating wider anti-social behaviours, with studies referencing diminishing rates of aggression and ‘violent content’ consumption across different sample groups. However, few pedagogical resources focus on how to specifically tackle extremism and radicalisation online. Similarly, there is currently little data enumerating the impact of these schemes on terrorist recruitment. A cluster of programmes are referenced in the literature, such as:


‘Don’t Be a Puppet’, an FBI-led intervention developed in partnership with ‘community leaders’ with the aim of educating young Americans on the ‘deceptive realities of violent extremism’. Conducted through a series of online quizzes and videos, Suzanne Waldman and Simona Verga cite the project as an example of ‘what holds promise’, although they concede that there is little evaluation data to substantiate such claims. The campaign has also been ‘broadly criticized for perpetuating stereotypes’ and ‘breeding suspicion over potentially innocuous activities’.

Methodological problems also afflict other case studies – including ‘Twin Track’ in the UK and the 2011 Royal Canadian Mounted Police (RCMP) project ‘Youth Online and at Risk’ – with assessments describing the guidance material itself rather than analysing how this content was consumed or whether it generated lasting results.

However, more robust findings are provided by Louis Reynolds and Ralph Scott in their analysis of ‘digital citizenship’ courses delivered to 165 students across four schools in the UK, with a post-pilot evaluation revealing statistically significant improvements in participants’ abilities to ‘think critically about online material’. Despite the small sample size and lack of randomisation, these results show promise and affirmed the value of ‘concrete skills development rather than general ideological messaging’. Other studies emphasise the merits of ‘value complexity’ in helping to dismantle the dichotomies and ‘othering’ processes embedded in VE content. Crucially, the use of dialogue-based rather than didactic approaches encouraged dynamic learning, reflecting the conclusions of other pilot studies that interrogate the utility of ‘flexible’ methods for teaching ‘e-safety’. This is complemented by encouraging results from even small-scale programmes building the capacity of teachers and delivering ‘digital resilience’

229. Davies et al., ‘Toward a Framework Understanding of Online Programs for Countering Violent Extremism’.
230. Waldman and Verga, ‘Countering Violent Extremism on Social Media’.
232. ISD and RAN, ‘Counter Narratives and Alternative Narratives’.
235. Ibid.
236. Ibid., p. 37.
curricula based on ‘best practice’. However, as evaluations of similar projects in Indonesia qualify, improving the skills and awareness of audiences is beneficial, but the longer-term effect on preventing susceptibility to extremist narratives is often yet to be evidenced.

Consequently, with the right blend of design and content, critical thinking and media literacy courses can potentially have positive effects, even though it is unclear how far they directly address the problem of radicalisation and recruitment.

The Right Messenger in Conversations and One-to-One Engagement

The involvement and testimony of former extremists (so-called ‘formers’) and defectors in P/CVE campaigns is a widely shared recommendation in the literature, and seems to be supported by theory. Michael Pizzuto, for instance, suggests that any messenger should be endogenous to the audiences they engage, applying lived experience to tap the same cultural and dialectical nuances exploited by local recruiters. These may include former extremists who, alongside the victims of terrorism, can de-glamourise violent groups and offer a more engaging set of emotive stories to cultivating trust and credibility. Yet, despite the significant material produced in conflict studies and analysis of ‘deradicalisation’, there is surprisingly little research assessing the role of former extremists or defectors in preventive work specifically. The few programmatic evaluations that are available appear to support the proposition, with interventions such as ‘ExitUSA’ – a messaging project launched by Life After Hate to tackle white supremacism – presenting indicative examples that former extremists can have resonance. Eight individuals requested additional support and information following exposure to the campaign, citing the credibility of the former extremists included in the ExitUSA team as part of their motivation, with Tanya Silverman and colleagues describing such feedback as ‘perhaps the most direct evidence possible of counter-narratives ... having impact’. Crucially,

239. Reynolds and Parker, ‘Digital Resilience’.
241. Alava, Fraa-Meigs and Hassan, Youth and Violent Extremism on Social Media.
243. Examples include outreach by the French Association of Victims of Terrorism, described in Jackson et al., ‘Practical Terrorism Prevention: Appendices’.
245. Significant case studies include Northern Ireland, Serbia, Lebanon, Burundi and Colombia.
these conclusions seem relatively cross-cutting: Abdullah X’s ‘appeal’ is partially ascribed to the previous extremist leanings of its founder,249 and projects like ‘EXIT White Power’ emphasise the value of stories from far-right defectors.250

However, as this paper highlights, there are limits to how far such metrics can serve as a proxy for impact-level data, and questions remain over how this material should be used and transmitted to maximise its efficacy. Given the scarcity of former extremists available for P/CVE programming, some argue that their experiences should be filmed, vetted, translated and globally distributed to expedite scalability and ensure a consistent narrative. The impact of testimonies were collated as part of ‘Break the ISIS Brand – the ISIS Defectors Interviews Project’ and tested on various focus groups, one of which included a former Islamic State emir, Abu Islam, now in Iraqi custody.251 Descriptive accounts of these sessions emphasise their emotive resonance: Islam looked down ‘apparently ashamed’, and the authors claim audiences across the Balkans, Central Asia, Western Europe and the Middle East generally accepted the videos’ authenticity and content.252 Anecdotal examples of effectiveness are also cited, such as the apparent dissuasion of a minor ‘hell-bent on going to Raqqa’.253 Nevertheless, conclusions need to be caveated by the composition of some focus groups: social workers, teachers and P/CVE practitioners may have valuable insights, but they are not necessarily representative of the target audience. In a more detailed analysis of tests with American colleagues and students, Speckhard and colleagues state they ‘did not ... attempt to reach or replicate the vulnerable populations that ISIS aims at’.254 Similarly, discussions were convened with Somali-American youths in San Diego, a context with historical experience of VE recruitment, but the sample itself was described as small and ‘unrepresentative’, and it is unclear how far the cohort actually reflected the demographics of those susceptible to Islamic


State outreach. Consequently, the films’ impact on ‘at-risk’ individuals remains somewhat ambiguous. Likewise, the brevity of tests limited any tracking of longer-term reactions, making it difficult to gauge the sustainability of any positive outcomes – a problem only accentuated by the lack of knowledge about who is accessing material online.

In contrast, ISD’s ‘One to One’ intervention sought to test the viability of directly contacting individuals expressing extremist sentiments online, engaging with 154 Facebook profiles displaying either ‘Islamist’ or far-right proclivities. Due to their experimental disposition, these activities were designed as an introductory approach to dissuasion rather than an attempt to completely deconstruct the ideological sympathies of participants. While the project’s indicators were output-oriented and the project was too short to ‘effectively measure any long term shifts in belief system or behavior’, the findings were nevertheless indicative. The ‘highest responses’ – a threshold of five or more messages – were generated from a combination of sentimentality, casual prompts and personal stories from former extremists, suggesting the impact of these messengers when working through interactive interventions.

Given the limitations associated with small sample sizes, ISD subsequently scaled up its approach with a follow-up project, ‘Counter Conversations’, involving over 800 selected candidates. The effectiveness of ‘formers’ appeared to be more mixed in this second iteration: they were most likely to prompt an initial response but delivered the ‘fewest number of conversations’ due to competing professional responsibilities. In contrast, other ‘frontline’ stakeholders such as survivors of extremist violence and professional counsellors were more successful at maintaining engagement and facilitating conversation respectively. This speaks to the wider issue of trust between the intervener and beneficiary when developing a dialogue. Marina Tapley and Gordon Clubb question the generally assumed credibility and suitability of former extremists in establishing these relationships, highlighting the contextual specificity and conditionality of

256. Ross Frennett and Moli Dow, ‘One to One Online Interventions: A Pilot CVE Methodology’, ISD and Curtin University, September 2015.
258. Frennett and Dow, ‘One to One Online Interventions’, p. 18.
259. Ibid.
261. Ibid.
262. Braddock and Morrison, ‘Cultivating Trust and Perceptions of Source Credibility in Online Counternarratives Intended to Reduce Support for Terrorism’; Davey, Birdwell and Skellett, ‘Counter Conversations’; Leuprecht et al., ‘Containing the Narrative’.
defectors’ legitimacy, particularly ‘within the radical milieu they left’.\(^{263}\) Their experiences are usually an essential ingredient in instigating conversation, but they may not have the specialist psychosocial or psychological skills to sustain or leverage outreach with vulnerable individuals. Insincere or imitative positions are also a risk if former extremists are influenced by, or perceived to be influenced by, ‘pressures from society or organisations they are involved in’.\(^{264}\) To expand and improve this coverage, former extremists may therefore require a significant investment in capacity-building, support infrastructure and pastoral care – services beyond the budgetary scope of many counter-messaging programmes.

More broadly, these findings suggest the credibility of messengers – regardless of their background – is contingent on a wider set of factors, demonstrating the importance of cultivating relationships and support structures that extend beyond the content of any singular message. Under these circumstances, Ross Frenett and Moli Dow suggest such content does not have to be bespoke as its currency is cumulative and derives from a sustained dialogue over time.\(^{265}\) Of the 800 selected candidates included in ‘Counter Conversations’, 569 were contacted, 112 responded to online outreach and 76 engaged in ‘sustained’ dialogue, of whom eight showed indicators of ‘potential positive impact’.\(^{266}\) This success rate may seem low in absolute terms, but it is high relative to other counter-messaging campaigns with comparable data, a reflection of the incipient baselines characterising much of the prevention space. The authors also acknowledge the difficulties of verifying the campaign’s influence. Noting that two in five individuals claimed ‘the conversation may have changed their mind’, Jacob Davey, Jonathan Birdwell and Rebecca Skellett caution that recipients may continue to ‘engage with extremist material in private’.\(^{267}\)

Nevertheless, the study is valuable in revealing clear generational deviations in these online dialogues and can be considered potentially effective, with one in 10 sustained engagements explicitly containing indicators of positive impact.\(^{268}\) Crucially, the traction generated by the pilot programmes derived from their dynamism and use of direct, private conversations as a mechanism for establishing relationships, trust-building and a wider process of socialisation.\(^{269}\) As Waldman and Verga summarise, combating or ridiculing extreme narratives ‘can help to dissuade young people from becoming involved in those groups, but only when opportunities are also available for them to have two-way conversations with frontline workers’.\(^{270}\)

---

264. Ibid., p. 6.
265. Frenett and Dow, ‘One to One Online Interventions’.
266. Davey, Birdwell and Skellett, ‘Counter Conversations’.
267. Ibid., p. 18.
268. Ibid.
269. Ibid.
The process of sequencing CVE content with dialogue and personalised engagement is therefore paramount, reiterating similar conclusions drawn from the fields of development and peacebuilding. Reviewing CAPE’s online discussion forum, Voogt similarly argues that respect, empathy, debate and the encouragement of critical thinking are ‘considerably enabled’ by one-to-one processes, in large part because implementers can speak to the specific experience, motivations and interests of vulnerable individuals.\textsuperscript{271} In this context, evidence underscores the need to help recipients reach conclusions rather than imposing paternalistic or combative prescriptions.\textsuperscript{272} While delivery appears to be most optimal when facilitated ‘face-to-face’,\textsuperscript{273} virtual interventions can also draw on ‘best practice’ from offline mediation such as ‘empathetic listening, multi-polar discussions, identifying shared values and common ground, and appealing to emotion’.\textsuperscript{274} However, there was little indication or analysis of efforts to synchronise on- and offline engagement in the literature collated for this review, making it difficult to compete with the ‘immersive … digital and physical connections’ offered by VE groups that anchors much of their recruitment processes.\textsuperscript{275} Thirty-seven of the studies addressed both mediums, but rarely analysed how they can intersect and complement one another, creating a significant gap in relation to an otherwise promising, if tentative, approach to P/CVE communications.

Positive Alternatives Within Integrated Interventions

Alternative narratives, emphasising ‘what we are for rather than what we are against’, are often hailed as more positive interventions than direct rebuttals of extremist propaganda.\textsuperscript{276} William D Casebeer and James A Russell argue terrorism can be more effectively countered by developing a ‘better story’ to replace VE narratives.\textsuperscript{277} As such, studies recommend leveraging ‘positive’ messages about social values, tolerance, ‘freedom and democracy’ to supplant solutions proposed in extremist discourse.\textsuperscript{278} This affords opportunities to be proactive and creative, encouraging the ‘rediscovery of cultural production’ that can use the ‘affective resources of story-telling’, rather than simply appealing to ‘reason or self-interest’.\textsuperscript{279} Crucially, it also creates scope for broadening and integrating P/CVE messaging with ‘other priorities’ of vulnerable individuals and wider communities, helping to generate ‘openings

\textsuperscript{271.} Voogt, ‘Countering Far-Right Recruitment Online’.
\textsuperscript{272.} Ibid.
\textsuperscript{273.} Richardson, ‘Fighting Fire With Fire’; Bilazarian, ‘Countering Violent Extremist Narratives Online’.
\textsuperscript{274.} Davey, Tuck and Amarasingam, ‘An Imprecise Science’, p. 22.
\textsuperscript{275.} Bilazarian, ‘Countering Violent Extremist Narratives Online’, p. 62.
\textsuperscript{278.} Beutel et al., ‘Guiding Principles for Countering and Displacing Extremist Narratives’.
\textsuperscript{279.} Glazzard, ‘Losing the Plot’, p. 16.
and new partnerships around CVE’. However, Talene Bilazarian draws a distinction between constructive synergy and the introduction of covert counterterrorism efforts under the rubric of ‘cohesion-oriented programming’.

‘Burka Avenger’ is a prominent case study in the literature. Broadcast as a mainstream TV programme across South Asia, the award-winning animation series features Jiya, a young heroine fighting for ‘peace, justice and education for all’. As the ‘most watched children’s show in Pakistan’, the project has cultivated considerable cultural capital that is reinvested in the promotion of ‘social awareness’ campaigns. Produced ‘by Pakistanis for Pakistanis’, Abbas Barzegar, Shawn Powers and Nagham El Karhili suggest that the Burka Avenger’s ‘artistic curators’ are able to ‘seamlessly weave powerful and diverse social forces into a coherent and resilient whole’, leveraging a local vernacular and aesthetic sensibilities ‘grounded in Pakistani national identity’. The elevation of ‘positive identities’ appears to motivate ‘collective responsibility’ and reproduce the mechanisms of ‘hero lionising’ so effectively exploited by groups like Al-Qa’ida and the Islamic State. Comic characters – such as ‘The 99’, a team of crimefighters (described as ‘inspired by Islam’ rather than ‘Islamic’), and Suleiman Bakhit’s ‘Hero Factor’ – subscribe to a similar logic, drawing on traditional religious concepts or cultural mythologies to develop appealing stories and role models.

Other examples include a nine-month pilot programme in Afghanistan which disseminated audio and video messages through memory cards on mobile phones, leveraging content designed by local teams using ‘recognisable images, themes, locations and language’. Little evidence links the initiative to a drop in local recruitment or detectable changes in the attitudes of recipients or wider public opinion, but it coincided with a 50% decrease in sales of extremist content across participating vendors and a concurrent rise in consumer demand for positive narratives, indicating the popularity of these alternative broadcasts – focused largely
on peacebuilding and social reconciliation – compared to those messages directly countering or attacking the Taliban.\textsuperscript{290}

However, any conclusions should be cautious and caveated. Rather than explicitly addressing violent discourse, such interventions generally promote constructive behaviours including ‘civic engagement’ that can have secondary P/CVE-specific outcomes, but create convoluted causal chains that can make it challenging to accurately gauge efficacy.\textsuperscript{291} These accounts are anecdotal, and widespread recommendations prescribing alternative narratives in the literature are not commensurate with the sparsity of available empirical evidence. While CSOs and private-sector stakeholders express an ‘overwhelming’ preference for alternative narratives\textsuperscript{292} – and both observational and theoretical analysis\textsuperscript{293} suggests they function as better vehicles for navigating sensitive issues and building resilience – the data remains elusive, particularly at a programmatic level.

It is also unclear how these narratives engage relevant target audiences.\textsuperscript{294} ‘Burka Avenger’ undoubtedly attracts huge ratings, but its message and underlying strength lies in a mainstream appeal, making it unclear whether those already exploring radical ideas or flirting with extreme ideologies would actually watch such content, and how they respond if they do. As with many analogous initiatives, little demographic analysis is accessible to resolve these questions, leaving the literature reliant on conceptual propositions. J M Berger, for instance, posits ‘even in a best-case scenario, positive messaging must seek the lowest common denominator’, resulting in vague value statements that are either ‘fatally watered down or hobbled by legitimate complaints of hypocrisy’.\textsuperscript{295} In contrast, Barzegar, Powers and El Karhili argue that ‘alternative narratives are equally if not more useful in reaching so-called fence sitters, as ... they are more capable of capturing audience interests and directing at-risk audiences in a more productive direction’.\textsuperscript{296} However, neither the authors nor the paper they cite provides empirical evidence to verify their assertions, and the latter largely focuses on disengagement rather than preventive processes, which are qualitatively discrete – if sometimes overlapping – issues.\textsuperscript{297}

\textsuperscript{292} Barzegar, Powers and El Karhili, ‘Civic Approaches to Confronting Violent Extremism’.
\textsuperscript{293} Schmid, ‘Al-Qaeda’s “Single Narrative” and Attempts to Develop Counter Narratives’.
\textsuperscript{294} Speckhard et al., ‘Mounting a Facebook Brand Awareness and Safety Ad Campaign to Break the ISIS Brand in Iraq’.
\textsuperscript{295} Berger, ‘Making CVE Work’, p. 5.
\textsuperscript{296} Barzegar, Powers and El Karhili, ‘Civic Approaches to Confronting Violent Extremism’, p. 28.
\textsuperscript{297} ISD and RAN, ‘Counter Narratives and Alternative Narratives’, RAN Issue Paper, October 2015, p. 5. The paper suggests counter and alternative narratives should be central to P/CVE, but offers little data beyond references to studies on disengagement or descriptions of P/CVE programmes to demonstrate the propensity of ‘positive’ content to better engage ‘fence-sitters’.
To entice and sustainably engage with an audience, studies emphasise the need to both acknowledge ‘the kernel of truth’ that often makes extremist propaganda so compelling and provide strategies for replacing the ‘perceived benefits of radicalisation’. As with counternarratives, discourse alone appears to be insufficient for alleviating the grievances, anxieties and interests driving recruitment. Consequentially, alternative narratives must not only be coupled with supplementary initiatives that offer viable pathways for achieving the positive approaches and lifestyles they advocate, but should also be realistic in their aspirations. Failure to do so may precipitate ‘say–do gaps’ – an incongruency between the ‘words’ and ‘actions’ of messengers – that undermine the credibility of such content or generate disparities between ‘positive alternatives’ proposed in messages and the tangible compensation practitioners can actually deliver, frustrating the expectations of their audience.

Interventions appear to be potentially effective when they are able to close such gaps and leverage communication strategically. Daniel P Aldrich’s evaluation of USAID’s ‘peace and tolerance’ radio programmes – conducted under a ‘CVE policy framework’ across Mali, Chad and Niger – demonstrates these dynamics as recipient communities also benefiting from concurrent educational, vocational and capacity-building projects were found to listen more regularly to the broadcasts, resulting in ‘measurably different behaviours and norms than their less tuned-in counter-parts’. Regression analysis further identified strong positive correlations between ‘higher levels of radio listening’, support for ‘working with the West to counter terrorism’ and ‘greater participation in local decision-making’. While these findings reveal the value of the transmissions themselves – raising public awareness, bypassing information bottlenecks, and creating ‘generative loops’ – they also show how multi-vectored approaches can amplify results. To qualify, these changes were restricted to lower-level cognitive and behavioural functions – for example, increasing the likelihood of continuing to tune in for ‘peace radio’ broadcasts and becoming more civically engaged – and had ‘little measurable impact’ on higher-level cultural norms or abstract beliefs, such as ‘whether the US [is] fighting Islam or terrorism’. It was likewise difficult to segregate intervention-induced impact from many other confounding variables, and the project was therefore deemed ‘potentially effective’.

298. Ibid.
299. Davies et al., ‘Toward a Framework Understanding of Online Programs for Countering Violent Extremism’.
300. Aldrich, ‘First Steps Towards Hearts and Minds?’; Waldman and Verga, ‘Countering Violent Extremism on Social Media’.
303. Ibid.
305. Aldrich, ‘Radio as the Voice of God’.
306. Ibid.
rather than ‘effective’ by this review. Nevertheless, Aldrich’s evaluation enumerates the benefits of complementing a positive communications campaign with a broader set of ‘engagement’ activities, highlighting statistically evident outcomes of peace messaging when blended with a tranche of long-term local development projects.\(^{307}\)

Reflecting on the mobile-based messaging campaign in Afghanistan, Casey Johnson reinforces these conclusions: ‘alternative narrative campaigns, no matter how locally tailored and creatively disseminated, will likely have a strictly limited impact if conducted in isolation of other programming’.\(^{308}\) Eran Fraenkel, Emrys Schoemaker and Sheldon Himelfarb similarly emphasise the importance of ‘face-to-face community outreach activities that provide ... an opportunity to put into practice ideas ... that the media have brought to their attention’;\(^ {309}\) and Ferguson draws analogies with previously effective multi-dimensional communications campaigns for HIV/AIDS that levied ‘medical provisions, practical strategies, community engagement’ and national legislation.\(^ {310}\) This speaks to the importance of multi-faceted projects that employ communication techniques holistically and strategically, rather than relying on a specific set of narrative-based activities. In such a context, positive alternatives, advocacy and stories can potentially both mitigate the damaging externalities produced by directly refuting extremist content, and supplement more convincing forms of outreach and engagement.

Prioritising Process

Although there are few examples in the literature, projects that build core skills, strengthen the agency of participants and encourage co-creativity have presented potentially effective results, especially when they are nested in wider trust- and relationship-building exercises. For example, film and theatre workshops have found traction in Afghanistan, providing vehicles for community dialogue, self-expression and trust-building. To address historic polarisation between madrassas and public schools, students were invited for a series of capacity-building activities focusing on editorial design, documentary production\(^ {311}\) or drama courses that encouraged pupils, teachers and parents to debate local social issues and use this as a basis for subsequent performances.\(^ {312}\) These courses provided mechanisms for navigating sensitive topics in an indirect way, facilitating dialogue and incentivising inter-communal collaboration. In contrast to other projects that grappled with the ethics of using theatrical productions to

\(^{307}\) Ibid.

\(^{308}\) Johnson, ‘Alternative Methods of Alternative Messaging’.


\(^{311}\) Sayara Research and USIP, ‘USIP Peaceful Elections and CVE Programming’.

\(^{312}\) Ibid.
convey counter-messaging, the value of these initiatives was not just the medium or final output, but the process it offered for learning and discussion.

Several shortcomings were nevertheless identified in relation to sustainability and longer-term impact. These schemes can create space for cathartic expression, but any benefits may be eclipsed by wider structural problems, especially if they do not receive follow-up or supplementary programming. Evaluations of similar youth empowerment workshops also flag the risk of a wider syllabus, encompassing skills-building, leadership and vocational training distracting from the substance of P/CVE programming. Moreover, it is a relatively insulated experience, where attention is confined to a small cohort of participants. Larger audiences can consume messages conveyed through theatre performances or films but cannot engage in the interactive process itself.

However, the holistic nature of Afghan interventions shows potential given the formation of ‘genuine friendships’, local efforts to continue joint projects and a greater sense of communal responsibility reported by students – shifts attributed to technical proficiencies developed through the workshops that have helped in their ‘personal careers’ and ‘self-awareness’. These ‘transferable skills’ also feed into broader efforts to strengthen citizen journalism – a type of programme that seems to have promise in peacebuilding and conflict resolution. In contrast to state-sponsored news outlets, Ferguson emphasises the value of independent, grassroots media groups in empowering marginalised voices and fostering ‘robust public debate’. Conflating these approaches too far with P/CVE runs the risk of securitisation, but they do have complementarity that can be leveraged, particularly if they are sequenced effectively. Unfortunately, such interventions remain ‘under-developed and under-resourced’, exposing a significant gap in the literature.

Similar mechanics have been taken further in initiatives like STRIVE (Strengthening Resilience against Violent Extremism), which developed a ‘preventive communications’ model to ‘strengthen relationships and build trust’ through various interventions – including a series of discussion-based activities in Kenya – to help at-risk individuals process and share their

314. Sayara Research and USIP, ‘USIP Peaceful Elections and CVE Programming’.
316. Sayara Research and USIP, ‘USIP Peaceful Elections and CVE Programming’.
317. Ibid.
320. STRIVE is an EU-funded project led by RUSI aimed at strengthening resilience to violent extremism in Kenya and Afghanistan.
In contrast to the orthodoxies of strategic communications, which interprets narrative as an ‘instrument for influencing others through its deliberate dissemination’, this approach considered ‘story as a medium’, allowing participants to examine their attitudes and emotions, and articulate issues relating to ‘experience, identity and relationship’. Consequentially, the underlying theory of preventive communications and other process-based projects reflects current thinking in development communications, emphasising the importance of an interactive journey and the centrality of relationships in outreach and engagement.

However, independent evaluations are ongoing, and it is too early to gauge any effect (or lack thereof). The replicability of STRIVE’s template may also be limited as it relied on a concurrent multi-year mentorship programme that helped strengthen bonds between practitioners, mentors and mentees, and offered a granulated understanding of each participant’s personality and frustrations. The introduction of any additional modules or exercises to therapeutically grounded interventions are contingent on trust, alluding to long timeframes and high costs often beyond the capabilities of smaller P/CVE programmes.

Such issues also reflect a broader problem of volume given the labour intensity of maintaining highly personalised activities, and many of these strategies do not lend themselves to economies of scale. This is a perennial problem, and Peter R Neumann argues that ‘even if we found the perfect message, the perfect messenger … it would still be a drop in the ocean’. Yet, there is an incongruous lack of investigation into how such activities can be systematised and scaled up, particularly those that integrate tailored interactions, participatory learning and one-to-one support structures. Until these gaps are addressed, many interventions will struggle to sustain and build on the progress they have achieved.

---

322. Freear and Glazzard, ‘Preventive Communication’.
326. Szmania and Fincher, ‘Countering Violent Extremism Online and Offline’; Davey, Birdwell and Skellett, ‘Counter Conversations’. 
Conclusions

This paper seeks to collate and analyse the available evidence underpinning P/CVE communication-based interventions; interrogating the literature to identify problems, gaps and ambiguities, and exploring avenues that appear to hold more promise. While the findings are invariably partial and subject to critique, they highlight important limitations that need to be recognised and opportunities that should be leveraged.

Finding 1: A vast evidence base with limited depth and breadth.

Available evidence for P/CVE communications is largely inconclusive, in part due to the paucity of data and a general reliance on output-level indicators and ‘vanity metrics’. There is scant assessment of outcomes or impact, and few longitudinal studies to understand whether interventions resonate and how any positive effects can be sustained over time. Despite its volume, the literature also displays limited breadth, frequently recycling a relatively small cluster of case studies and predominantly focusing on counternarratives rather than the wider suite of activities included under ‘strategic communications’. This leaves research on P/CVE communications both oversaturated and underdeveloped.

Important gaps need to be addressed, including: more analysis of existing interventions; a wider focus on ‘holistic’ communication-based activities; a greater interrogation of how contextual and social dynamics frame P/CVE outreach; a broader appraisal of theoretical, empirical and operational lessons from related disciplines and fields; and further investigation into how both ‘natural world content’ and youth leadership can be harnessed and supported.

Finding 2: Fragmented theory and problematic assumptions are acknowledged but need to be addressed.

Assumptions and outdated theoretical frameworks pervade P/CVE communications – most prominently, the claim that ‘violent words lead to violent deeds’, a proposition based on the presumed linearity between consuming extremist content and radicalisation and recruitment. Consequently, attention has largely focused on countering VE material and prioritising the ‘message’, overlooking the variables that actually foster appetite for extremist material in the first place, and neglecting a complex matrix of personal motivations, structural drivers and affective bonds (online and offline) forged through recruitment.

Crucially, these are not novel findings: the assumptions and theoretical constraints identified in this review are widely acknowledged in the literature. However, they continue to permeate the programmatic logic of many counter-messaging interventions. Such collective dissonance has led to

327. Ferguson, ‘Countering Violent Extremism Through Media and Communication Strategies’.
a degree of atrophy across P/CVE communications even as individual projects display creativity and experimentation, reflecting a need to not only critique these normative conventions in theory but to also respond and replace them in practice. Such a shift requires a wider reappraisal of the purpose and scope of communication-based programming rather than the routine prescription of technical improvements that dominate the literature.

Similarly, interventions should engage with the mechanics actually underpinning VE recruitment processes and build on, rather than just relearn, the many lessons identified by academics and practitioners.

**Finding 3: The ‘right’ messenger is transient, context-dependent and necessary but not sufficient.**

The credibility of government-fronted discourse remains spurious in the context of P/CVE communications, and the importance of an authentic messenger as part of a successful counter-messaging ‘formula’ is widely acknowledged. However, there is less interrogation of the agency, capacity and palatability of CSOs across different social milieus, or empirical analysis assessing the role of other stakeholders, such as former extremists, in designing and delivering content. Understanding and tracking where these actors – alongside religious authorities, communal leaders and other traditional P/CVE interlocutors – feature in the networks and social circuitries of vulnerable populations is essential as there is a risk that interventions may overlook the legitimacy exercised by informal authorities, friends and other unconventional influencers.

A corollary is the importance of identifying the best combinations and mode of programming to ensure each stakeholder is engaged effectively. For example, local stakeholders appear to show greater traction when they draw on the same epistemological and normative foundations as VE recruiters, but do so when discussing a range of issues rather than rebutting specific ideological claims. The experiences and profile of intervening ‘messengers’ are important in determining how these dynamics function, but the process of sequencing P/CVE content with dialogue, relationship-building and personalised engagement is paramount, reiterating similar conclusions drawn from the fields of development and peacebuilding.

Practitioners also need to accurately diagnose the limitations and necessary support systems, training and investment these messengers may require, particularly in the case of former extremists. Similarly, the risks of securitisation and perceived co-optation should be fully considered and mitigated: social networks are not static and will likely respond negatively if messengers are seen as a front for state propaganda.

**Finding 4: A good online message is rarely enough.**

The evidence base remains disproportionately focused on digital interventions – an understandable fixation given the quantity of malicious content on the internet and the accessibility violent extremists can exploit through social media and other Web 2.0 platforms. While a comprehensive review of recruitment dynamics was beyond the scope of this paper, studies suggest that exposure to online VE propaganda often tends to reinforce already made up minds. Mobilisation is instead
largely facilitated and accelerated by affective and relational dynamics both on and offline that transcend discursive content. In contrast, counter-messages are frequently circulated without wider support networks, diminishing the resonance of their message and leaving the social bonds created by radicalisation and recruitment processes unaddressed. They can also struggle to identify and engage target audiences and, when operating exclusively online, receive little indication of cognitive and behavioural change. Consequentially, these approaches are often framed by limited empirical evidence of effectiveness and significant theoretical shortfalls.

This does not negate the value of such interventions but emphasises the importance of synthesising on- and offline engagement in ways that reflect and out-compete the creation of ‘immersive … digital and physical connections’ by VE groups. Embedding outreach within wider networked approaches, leveraging interpersonal dialogue and building mechanisms for follow-up and long-term engagement, alongside proactive offline activities, are therefore essential. More research and experimentation is needed to understand how to best synchronise such efforts given the current lack of evaluation material. Strengthening the digital literacy, awareness and critical thinking skills of individuals interacting with online content could also be beneficial as a supplementary approach.

Finding 5: Integrated interventions are essential.

More broadly, evidence indicates the need to move beyond a rather reductive preoccupation with counternarratives and alternative narratives, especially those delivered in isolation. This is not to dismiss their role completely, as studies highlight their utility in various contexts and under certain conditions, but greater investment is needed in programmes that leverage ‘communication’ strategically and holistically. Supplementing messages within a wider set of activities can potentially help resolve say–do gaps, build credibility and create viable alternatives. This not only helps strengthen trust in – and the legitimacy of – those broadcasting content, but also markets a more comprehensive, tangible and attractive package commensurate with the benefits offered by VE groups.

Finding 6: The process of highly personalised engagement can itself be valuable, although difficult to scale up.

Trust, socialisation and relationship-building are critical pillars of both VE and P/CVE. Pilot schemes that facilitate and encourage personal connections through, for example, one-to-one dialogue and conversation demonstrate a degree of traction. Likewise, interactive models that establish wider support systems may potentially be effective, especially when they are process-oriented. However, there are questions over the feasibility of scaling up such interventions when they are so resource- and labour-intensive, requiring both time and personalised engagement. Investing in opportunities for expanding and sustaining these approaches is therefore critical, as is the need to understand whether any positive benefits ‘stick’ or diminish once the process has concluded.

About the Author

Michael Jones is a Research Fellow in the Terrorism and Conflict team at RUSI, examining political violence, conflict economies and the convergence of violent extremism and insurgent militancy in East and sub-Saharan Africa. He has led investigative fieldwork across various countries including Sudan, Kenya and Lebanon.
Annex I: Bibliography

Screened Studies Selected for the Review of this Thematic Paper


Alex Wilner and Brandon Rigato on behalf of the ‘60 Days of PVE’ group, ‘The 60 Days of PVE Campaign: Lessons on Organizing an Online, Peer-to-Peer, Counter-Radicalization Program’, Journal for Deradicalization (No. 12, 2017), pp. 227–68.


Angela Trethewey, Steven R Corman and Bud Goodall, ‘Out of Their Heads and Into Their Conversation: Countering Extremist Ideology’, Consortium for Strategic Communication, Arizona State University, September 2009.


Anne Speckhard et al., ‘Mounting a Facebook Brand Awareness and Safety Ad Campaign to Break the ISIS Brand in Iraq’, *Perspectives on Terrorism* (Vol. 12, No. 3, June 2018), pp. 50–66.


Garth Davies et al., ‘Toward a Framework Understanding of Online Programs for Countering Violent Extremism’, *Journal for Deradicalization* (No. 6, Spring 2016), pp. 51–86.


H L Goodall Jr et al., ‘Rhetorical Charms: The Promise and Pitfalls of Humor and Ridicule as Strategies to Counter Extremist Narratives’, *Perspectives on Terrorism* (Vol. 6, No. 1, March 2012), pp. 70–79.


Jamie Bartlett and Alex Krasodomski-Jones, ‘Counter-Speech Examining Content That Challenges Extremism Online’, Demos, October 2015.


Naomi Theuri, ‘Gender and Contextual Perspective in Countering Violent Extremism (CVE): Examining Inclusion of Women and Contextual Factors in Online Approaches to CVE’, Degree Project, Department of Criminology, Malmö University, 2017.


Roslyn Richardson, ‘Fighting Fire with Fire: Target Audience Responses to Online Anti-Violence Campaigns’, Australian Strategic Policy Institute, December 2013.

Ross Frenett and Moli Dow, ‘One to One Online Interventions: A Pilot CVE Methodology’, ISD and Curtin University, September 2015.


**Supportive Literature Referenced in this Paper**


Annex II: Research Methodology

In January 2018, the Norwegian government commissioned RUSI to lead the Prevention Project, which ran for over two years. The project aims to improve the knowledge base for preventing and countering violent extremist programming. Facing stark conceptual and methodological challenges (outlined in detail below), preventive interventions have generally relied on assumption-based logics with little empirical grounding, exposing the field to a range of theoretical, practical and ethical problems.

By attempting to answer the research question ‘what can work and what has not worked in preventing/countering violent extremism (P/CVE)?’, the Prevention Project addresses some of these shortfalls, synthesising academic papers, evaluations, policy briefs and internal documents to understand what evidence, if any, exists for the ‘successful’ or effective application of such activities. This process condensed key findings from the literature and interrogated the basis of these findings to critically assess the substance and limitations of the source material with the aim of understanding the effectiveness (or not) of the intervention approaches described in the literature.

The approach to this review involved: 1) identification of search terms and criteria for inclusion and exclusion; 2) identification of potential sources; 3) collection of material related to P/CVE interventions using key search terms; 4) identification of additional material through snowballing; 5) removal of any material that was not relevant to this study and grouping of collected material into the relevant ‘thematic’ categories; 6) scoring of these studies according to their quality and assigning a related grading (high, medium or low quality); and 7) analysis of the documents to diagnose common assumptions or theories of change underpinning each thematic intervention, the validity of these assumptions and the effectiveness (or not) of the intervention described in the document.

From the outset, it is important to highlight that this was not a systematic literature review in the traditional sense. Systematic methods and principles were, however, adopted where possible to improve transparency, rigour and breadth, and to gauge the robustness of available evidence. In contrast to the natural sciences where this approach was pioneered, there is an ‘inherent contradiction’ between the information required to conduct a systematic review and

---

1. The project drew on previous work conducted with Eric Rosand and the similarly named ‘Prevention Project: Organising Against Violent Extremism’. The collaborative relationship with Eric continued for the duration of this project. For more information, see Organizing Against Violent Extremism, ‘About the Prevention Project’, <https://organizingagainstve.org/about-the-prevention-project/>, accessed 30 April 2020.
the structure, variance and content of social science studies. The reliance on non-positivist, qualitative methodologies which generally define these disciplines creates challenges: commensurate quality appraisal techniques lack consensus and remain relatively undeveloped. Systematic reviews have also struggled to adequately capture ‘less tangible, difficult to measure outcomes’, such as those in P/CVE, especially when they are nested in or intersect with wider processes and contextual dynamics. Greater flexibility was therefore necessary to accommodate these limitations, and this paper describes the methodological approach adopted for this project in full.

The Literary Landscape and its Limitations

P/CVE has been contested and critiqued on numerous fronts, from being overly reactive and externally imposed, to infringing on civil liberties, unfairly discriminating against ‘suspect communities’, and producing unintended outcomes and negative externalities. It has also been accused of lacking a coherent strategy and for being imbued with definitional and conceptual problems.

3. Ibid.
4. Ibid.
A Confused Vocabulary

P/CVE is generally considered to be a broad umbrella term to ‘categorise activities implemented by governmental and non-governmental actors seeking to prevent or mitigate violent extremism through non-coercive measures that are united by the objective of addressing the drivers of violent extremism’. However, linguistic ambiguities and confluations are widespread in the P/CVE space. This is in large part because many stakeholders tend to use ‘countering violent extremism’ (CVE) and ‘preventing violent extremism’ (PVE) interchangeably, arguing that there is little difference in objectives, mechanisms or actions between the two. Some development organisations, practitioners and scholars may opt for the PVE label to help distinguish upstream preventive approaches from any ‘security driven framework’, criticising CVE as a vehicle for ‘securitisising’ civic domains, such as healthcare, social work and education, and highlighting the term’s genesis in the US-led ‘Global War on Terror’. However, the lack of a consistent definition means it is not possible to draw comparisons between the relative benefits of preventing or countering approaches.

Even within the UN system there are significant discrepancies: for instance, the Security Council’s Counter-Terrorism Executive Directorate and the United Nations Office for Counter-Terrorism use the terms ‘CVE’ and ‘PVE’ respectively, despite sharing a relatively homogenous understanding of the steps necessary to diminish the threat of violent extremism (VE). Both agencies also occasionally conflate these appellations as P/CVE, exemplifying the inconsistency in the application of terminology.

This contestation extends to the adjunct processes of radicalisation and recruitment. The former has various definitions but is generally understood as the ‘social and psychological process of incrementally experienced commitment to extremist ideologies’. This is considered to be a fluid, non-linear and largely idiosyncratic process that affects people in different ways, and does not necessarily imply the adoption of violent behaviour. Instead, radicalisation involves a transition from ‘relatively mainstream beliefs’ to seeking some ‘drastic’ social and/or political change, which may or may not involve violence. Despite the tendency to frame radicalisation

---

10. Rosand et al., ‘A Roadmap to Progress’.
12. Ibid.
as a recognisable and consistent phenomenon, it is a concept that is often applied loosely to an eclectic mix of cases and situations.\(^{15}\)

In contrast, Edgar Jones describes recruitment as a ‘dynamic process by which a willing or unwilling individual is encouraged or dissuaded from joining a group; it involves a measure of assessment on both sides’.\(^{16}\) This is therefore distinct from, but may overlap with, the ‘belief modification’ associated with radicalisation.\(^{17}\)

**Conceptual Problems**

Crucially, P/CVE also faces constraints and ambiguities as VE ‘cannot be neatly packaged’\(^{18}\) due to its discrete iterations and drivers, leading to a myriad of potentially relevant intervention types, including: community debates on sensitive topics; media messaging; interfaith dialogues; empowerment programmes (particularly of women); training of government and security officials; and programmes aimed at individuals deemed to be ‘at risk’ of joining or being attracted to violent extremist groups. Consequently, ‘prevention’ risks become a ‘catch-all category’ that conflates with ‘well-established fields, such as development and poverty alleviation, governance and democratization, and education’.\(^{19}\) The mislabelling and ‘re-hatting’ of development interventions alongside the covert nature of many preventive activities accentuates this problem, making it difficult to systematically identify P/CVE programming in both theory and practice.

This is compounded by the amorphic nature of VE itself, a phenomenon that is difficult to clearly differentiate from a wider spectrum of violent action, from insurgencies to pogroms and local riots. The UN has notably failed to develop any universally recognised definition of either ‘violent extremism’ or ‘terrorism’,\(^{20}\) and delineations made in the literature are typically context-dependent and often contradictory, especially given the sensitivities and politicisation of such labels. Afghanistan, for instance, is considered an important arena for preventive interventions,\(^{21}\) but staple case studies in conflict analysis, such as Northern Ireland, Sri Lanka and Colombia rarely appear in the P/CVE discourse, despite all four appearing as comparative examples for assessing counterterrorism, disengagement and deradicalisation. This disjuncture

---

15. RUSI, ‘Countering Violent Extremism Curriculum’.
20. Rosand et al., ‘A Roadmap to Progress’.
exposes clear discursive, conceptual and theoretical problems with ‘violent extremism’ as a distinct analytical category due to its overlap with wider conflict ecologies.

The genealogy of P/CVE as a concept and a policy domain are also inextricably tied to ‘Islamist-based terrorism’ given its association with the ‘Global War on Terror’. It has since grown in both popularity and scope, integrating other manifestations of VE, such as white supremacism and residual strands of neo-fascism. Nevertheless, there continues to be a disproportionate focus on violent ‘jihadism’, meaning the true breadth of extremist militancy, replete with its numerous derivatives and sub-categories, is rarely represented in the literature.  

In such a confused context, the ‘public health model’ has become an increasingly prominent method for organising and reinterpreting P/CVE activity and agency, drawing on tested approaches for triaging ‘disease responses’ and healthcare. There are various iterations of this framework, but they generally distinguish between three levels of intervention: primary; secondary; and tertiary. Figure 1 demonstrates the authors’ approach to the model adopted for this research project.

---

22. This disparity appears to be less pronounced in the ‘deradicalisation’ literature, where there has been a prominent strand of academic and practical engagement with demobilising members of far-right groups.

23. There are numerous examples of the public health model framework. See, for instance, Jonathan Challgren et al., ‘Countering Violent Extremism: Applying the Public Health Model’, Center for Security Studies, Georgetown University, October 2016.

24. Some versions add a fourth level – ‘primordial’ prevention – at the base of the pyramid, meaning social and economic policies which affect health.
Figure 1: The Public Health Model for P/CVE

- **Primary**: Broad-based and community-focused prevention programmes addressing a range of social ills including, but not specifically focusing on, factors contributing to radicalisation and/or recruitment into VE.
- **Secondary**: P/CVE activities that either target populations/individuals identified as being ‘at risk’ or vulnerable to radicalisation and/or recruitment, or address individual incentives, enabling factors and structural motivators contributing to VE. This category has been expanded from the original model proposed by Jonathan Challgren and colleagues, described as activities focused towards ‘individuals and groups identified as at-risk for violent extremism’. The addition of interventions that include P/CVE objectives in their explicit or implicit theory of change and/or those addressing factors specifically contributing to recruitment and radicalisation helps reflect contextual and programmatic heterogeneity in what is a sprawling, largely ill-defined domain.
- **Tertiary**: Engaging individuals who have already joined terrorist groups or are identified as violent extremists, these activities typically include disengagement, deradicalisation, isolation and redirection, or counterterrorism.

This is not a perfect typology, especially given the porosity of its conceptual boundaries and potential inconsistencies when applied across heterogenous contexts, which introduces a degree of subjectivity when distinguishing between tiers. Nevertheless, the model is useful.

for reconfiguring an otherwise convoluted P/CVE sector, highlighting the goals, mechanisms and target audiences of various activities as they respond to different stages of radicalisation and recruitment, and demonstrating how they interact and synchronise with one another.

Problems in Data Collection and Quality

Stakeholders working in the P/CVE space have long described a general lack of good-quality data, especially in relation to monitoring and evaluation. For instance, the University of Maryland’s National Consortium for the Study of Terrorism and Responses to Terrorism found only five studies reporting outcome data assessing preventive programmes/interventions between 2005 and 2015, and other studies highlight both the limited availability and questionable quality of a large proportion of P/CVE content. This is the result of various methodological restrictions that are not unique to the P/CVE space but remain pronounced:

- **Problems of Attribution:** The programmatic logic of a preventive intervention or its ‘theory of change’ can often become incoherent if it extends too far upstream, as the pathway from delivery to impact of end-target groups is increasingly contorted or

---

26. This does not imply any linear relationship between different stages but simply reflects the intensity of cognitive and/or behavioural change within individuals during their own specific trajectory of radicalisation and/or recruitment.

27. Challgren et al., ‘Countering Violent Extremism’.

28. Caitlin Mastroe and Susan Szmania, ‘Surveying CVE Metrics in Prevention, Disengagement and Deradicalization Programs’, National Consortium for the Study of Terrorism and Responses to Terrorism, University of Maryland, March 2016.


Understanding and tracing these relationships within a litany of variables is difficult, especially when evaluators cannot disaggregate the specific impact of a project from other activities conducted in the same space, or segregate any effect from concurrent shifts in the wider milieu. This leaves attribution difficult to establish, with the lack of short, manageable causal chains making it challenging to exclude rival explanations for a specific trend or effect.\textsuperscript{32} Moreover, intended outcomes in P/CVE usually involve ‘nothing happening’, for example, the absence of radicalisation and recruitment. Assessing the mechanics of interventions is therefore problematic as any metric relies on an imperfect set of proxies to ‘prove a negative’, particularly as ethical constraints in complex and challenging contexts usually preclude any comparison between treatment and control groups.

- **Indicators of Success**: Given the diversity of focus areas, confused or contested models of radicalisation, and congruently vague policy objectives, it is hard to formulate indicators of success that relate concrete measures to impact on beneficiaries.\textsuperscript{33} Many expected outcomes in P/CVE involve ephemeral changes related to cognition, perception and opinion, which are challenging to track, especially with a paucity of secure baselines for comparison.

- **Operational Challenges**: Stakeholders are often reticent to divert resources away from core programming and there is little appetite on the part of local practitioners to publicise their ‘failures’ as this could compromise future funding opportunities.\textsuperscript{34} Similarly, evaluations are encumbered by the immaturity of preventive projects: many long-term interventions have not yet concluded, and completed programmes are frequently designed with short time horizons, limiting avenues for longer-term or longitudinal analyses. Information sharing also relies on a culture of transparency and receptivity, which is difficult to manage when data is sensitive, securitised or heavily regulated.\textsuperscript{35}

Consequently, monitoring and evaluation in the field of P/CVE tends to concentrate more on programmatic outputs to demonstrate the functionality and efficiency of individual activities. These results are usually difficult to generalise and offer little substantive assessment on the effectiveness of projects beyond superficial benchmarks that do not account for externalities or indirect and long-term impact.\textsuperscript{36} Where attempts are made to enumerate outcome-level findings, data is often ‘anecdotal and descriptive’, making inferences about effectiveness that are conjectural, ‘dependent on narrative interpretation’ and ‘difficult to validate’.\textsuperscript{37}

\begin{footnotes}
\item[32] Ibid.
\item[33] Ibid.
\item[34] Rosand et al., ‘A Roadmap to Progress’.
\item[36] Lindekilde, ‘Value for Money?’.
\item[37] Ibid.
\end{footnotes}
Given these limitations, it is therefore important that any enquiry into what can work and what has not worked in the P/CVE space establishes how robust the evidence base actually is, identifying not only what the literature claims but interrogating what these claims are based on.

**Methodological Approach**

As noted at the beginning of this paper, there were seven stages to the literature review. These are outlined in detail below.

1. **Search Terms and Criteria for Inclusion and Exclusion**

As part of the literature review for this project, the team designed a set of inclusion/exclusion criteria that would ensure adequate coverage in its data-collection:

**Table 2: Inclusion and Exclusion Criteria**

<table>
<thead>
<tr>
<th>Geographical Locations</th>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Language</th>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Other languages</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Conceptual Focus</th>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only P/CVE interventions aimed at the secondary level of the adapted public health model, defined as: 1) interventions that label themselves as PVE, CVE or P/CVE, counter-radicalisation, etc.; 2) interventions that identify factors of VE and how they will address these; and 3) interventions that identify ‘at-risk’ and ‘vulnerable’ populations or individuals.</td>
<td>Interventions that do not satisfy these criteria, primary and tertiary-level interventions (for example, deradicalisation, disengagement and reintegration).</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Types of VE</th>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>All types</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Publication Date</th>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005–present</td>
<td>Pre-2005</td>
<td>N/A</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Publication Format</th>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Peer-reviewed academic outputs, including journal articles, working papers, e-books and other online resources, and other academic outputs; 2) grey literature, including discussion papers, policy briefs, journalistic accounts, conference papers, good practice guidelines and toolkits; and 3) evaluations assessing impact, including independent and self-evaluations.</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

*Source: Table generated by authors based on the team’s inclusion/exclusion criteria.*
As noted in Table 1, only publications that focused on interventions falling within the secondary level of the authors’ adapted public health model were included. While there are overlaps with other tiers, the huge suite of activities included in primary-level programming, and their often-convoluted relationship with VE as a specific social ill, is beyond the scope of this project. Tertiary interventions engage those who are already violent extremists and subscribe to a distinct set of logics, mechanisms and processes. As a result, this category was also excluded to prioritise a focus on prevention work.

While inconsistencies in the labels of both radicalisation and recruitment have been highlighted, programmes were included in this review irrespective of their chosen definitions for one or both processes, as long as the programme itself aligned with secondary-level criteria enumerated in the public health model. This is largely because the Prevention Project sought to accurately interrogate the literature within its own self-defined parameters and was therefore forced to replicate any discrepancies it found when mapping the P/CVE ‘evidence base’.

2. Identification of Potential Sources

Having defined the inclusion/exclusion criteria, the team’s experience, contact networks and well-known P/CVE knowledge hubs were leveraged to map out sources for a multi-track data-collection process. As outlined below, these not only included ‘traditional peer review storage systems’ but also ‘alternative channels’ to ensure adequate coverage of grey literature and other content typically omitted from the conventional ‘information architecture’ characterising both P/CVE and the wider development space.38

- **Online search engines**, including JSTOR, Science Direct, Scopus, Web of Science, Google Scholar and British Library catalogues.
- **Official websites of international and regional donors**, such as the UN, the EU, the African Union, the Intergovernmental Authority on Development, the Global Counter-Terrorism Forum, the Global Community Engagement and Resilience Fund and various European, Middle Eastern, Asian and African governments, alongside the US and Canada.
- **Websites of key stakeholders, NGOs and practitioners**, such as the Institute for Strategic Dialogue, Mercy Corps, International Alert, Search for Common Ground, Overseas Development Institute, the British Council, CIVI.POL, the Global Center on Cooperative Security, and the Anti-Violent Extremism Network, among many others.

3. Collection of Material Related to P/CVE Interventions Using Key Search Terms

A list of ‘search terms’ was then developed, with the emphasis on P/CVE to avoid an overwhelming number of responses. As highlighted in the inclusion/exclusion criteria, the explicit inclusion of P/CVE terminology allowed a prioritisation of those studies that specifically focused on the issue of VE rather than wider development and peacebuilding issues.

Table 3: Search Terms

| Search Terms 1 | PVE, CVE, P/CVE, counter-radicalisation, prevent [prevention], ‘preventing violent extremism’, ‘countering violent extremism’ |
| Search Terms 2 | evaluate [evaluating/evaluate/evaluation], impact, evidence, review; effective [effective/effectiveness], ineffective [ineffective/ineffectiveness], challenges, success [successes/successful], failure [failed/failing] |
| Logical Operators | And/Or |

Source: Table generated by authors based on the team’s chosen search terms.

4. Identification of Additional Material Through Snowballing

This was supplemented with a series of forward and backward snowballing processes. Using the references and bibliographies of collected papers, any relevant studies omitted from the initial search were identified and several P/CVE experts were contacted for further direction and suggestions. Hand searches were subsequently conducted on Google to capture any remaining documents, particularly ‘non-academic’ articles, newly released studies and content on preventive work (either explicitly working with vulnerable individuals susceptible to recruitment and/or radicalisation or tackling any drivers/factors identified as contributing to VE) without clear labelling of these efforts as P/CVE interventions.

5. Removal of Any Material that was Not Relevant to this Study and Grouping of Collected Material into ‘Thematic’ Categories

These documents were individually screened by each team member to ensure the satisfaction of inclusion criteria. Any documents that did not meet the inclusion criteria were removed at this stage. The remaining documents were divided into the specific types of thematic intervention that were dictated by the reviewed literature: ‘women-focused interventions’; ‘religiously based mechanisms; ‘education’; ‘mentorship’; ‘P/CVE communications’; ‘youth empowerment’; ‘social cohesion/resilience’; ‘economic empowerment’; and ‘human rights and law enforcement’. In practice, many of these interventions are overlapping – for example, documents addressing mentorship programmes can also explore how critical thinking programmes are used in education. Therefore, certain studies overlapped between categories, especially those examining multiple or multifaceted programmes. Accordingly, these articles were scored once and integrated across the relevant thematic papers.
6. Scoring of These Studies According to Their Quality and Assigning a Related Grade (High, Medium, Low)

The articles were then classified through a rapid evidence assessment to score each paper’s ‘quality’. Quality was assessed according to a fixed set of criteria: conceptual framing, transparency, method, research design, internal validity, and cogency, replete with a series of sub-questions as detailed below.99

Table 4: Quality Scoring Criteria

<table>
<thead>
<tr>
<th>Scoring</th>
<th>Category</th>
<th>Sub-Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Conceptual Framing</td>
<td>• Does the study acknowledge existing research?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Does the study lay out assumptions and describe how they think about an issue?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Does the study pose a research question or outline a hypothesis?</td>
</tr>
<tr>
<td>3</td>
<td>Transparency</td>
<td>• What is the geography/context in which the study was conducted?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Does the study present or link to the raw data it analyses?</td>
</tr>
<tr>
<td>3</td>
<td>Method</td>
<td>• Does the study identify a research method?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Does the study demonstrate why the chosen design and method are well suited to the research question?</td>
</tr>
<tr>
<td>3</td>
<td>Research Design</td>
<td>• Does the study employ primary research methods?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Does the study employ secondary research methods?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Does the study rely exclusively on a theoretical or conceptual premise?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(As explained in DFID’s ‘How to Note’, ‘most studies (primary and secondary) include some discussion of theory, but some focus almost exclusively on the construction of new theories rather than generating, or synthesising empirical data’.)</td>
</tr>
<tr>
<td>3</td>
<td>Validity</td>
<td>• To what extent is the study internally valid for achieving its objectives?</td>
</tr>
<tr>
<td>3</td>
<td>Cogency</td>
<td>• Does the author ‘signpost’ the reader throughout?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• To what extent does the author consider the study’s limitations and/or alternative interpretations of the analysis?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Are the conclusions clearly based on the study’s results?</td>
</tr>
<tr>
<td>1</td>
<td>Independence</td>
<td>• Is the assessment conducted by an independent party (to those conducting the intervention itself)?</td>
</tr>
</tbody>
</table>

Source: Based on the ‘Principles of Quality’ from DFID’s ‘How to Note’ (p. 14) but adapted to reflect the scoring criteria for the ‘Prevention Project’.

Aside from the ‘independence’ category, which entailed a binary score of 0 or 1, the articles were assigned a value of 0 (absent) to 3 (strong) for each category. Team members swapped and re-scored samples of the documents to control for human bias, subjectivity and variation where possible. Once the articles were scored, the scores were aggregated and each paper was given a quality grading. Scores of 0–9 were graded as ‘low quality’; 10–14 were ‘moderate quality’; and 15–19 were considered ‘high quality’.

Two important aspects to this process need to be noted. First, quality was not an inclusion criterion in this study. Instead, the decision was deliberately taken to focus on quantity over quality in order to develop an evidence base. The quality grading was used during the analysis process to understand the weight and significance to ascribe to each paper’s findings and conclusions. Second, although quality was taken into account in the analytical process, the authors have refrained from associating (public) gradings to each reviewed study in the publication series out of respect for the work of other scholars in the field. It is also acknowledged that the grading system may have certain biases, as explained below.40

7. Analysis of the Documents in Order to Identify Common Assumptions, Assess the Validity of These Assumptions and the Effectiveness (or Not) of the Intervention Approach Described

Once the literature was graded, the documents were analysed to diagnose common assumptions or theories of change of each thematic intervention. The validity of these assumptions was subsequently explored using the evidence presented in the different papers. This includes an interrogation of the claims made in the articles – for example, were their claims substantiated by the data presented? Were any conclusions commensurate with the evidence presented in the study? What assumptions or conclusions were not verified?

During this interrogation, the research team assessed whether the assumptions underpinning the intervention(s) were valid and effective. This assessment was based on: the study’s own assessment of impact, if available; an analysis of the evidence or data presented to support this

40. Anyone interested in obtaining information on these gradings for educational or research purposes can contact the authors directly for more information.
assessment; and the quality grading of each paper. Each paper was then coded as ‘effective’, ‘potentially effective’, ‘mixed’, ‘ineffective’ or ‘inconclusive’:\footnote{41}

- Studies identifying a positive impact in relation to specific P/CVE objectives that could either be traced back to the contributions of a specific project, or causally attributed to an intervention, were regarded as ‘effective’.
- Studies that based conclusions on intermediate outcomes or anecdotal evidence of success were regarded as ‘potentially effective’.
- Studies that found that interventions produced both positive and negative results were categorised as ‘mixed’.
- Studies concluding that the intervention failed to produce the desired results were regarded as ‘ineffective’, while studies with an absence of any clear findings or those describing a project’s results as ambiguous were deemed ‘inconclusive’.

A tabulated summary of the team’s assessment of the evidence base for each thematic category, based on the aggregation of both ‘quality’ and ‘effectiveness’ assessments, are included in each thematic paper in this publication series.

There are nine thematic publications in this study as dictated by the literature gathered. These explore: ‘women-focused interventions’; ‘religiously based mechanisms’; ‘education’; ‘mentorship’; ‘P/CVE communications’; ‘youth empowerment’; ‘social cohesion/resilience’; ‘economic empowerment’; and ‘human rights and law enforcement’.

These are accompanied by two case studies exploring P/CVE in practice in Kenya and Lebanon. These countries were selected as areas where there has been a saturation of P/CVE activities and interest from a range of donors, including the Norwegian government. RUSI also has a strong foothold in Kenya given its office in Nairobi, which leads a P/CVE programme — STRIVE (Strengthening Resilience against Violent Extremism) II.\footnote{43} The two case studies will detail whether and how primary research fed into the results of the analysis exploring P/CVE interventions in practice in each country.

\footnote{41. Our definition of (in)effectiveness drew on OECD, ‘Evaluation Criteria’, <https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm>, accessed 15 March 2020. However, given that significant numbers of the reviewed studies were not evaluations, the categories of effectiveness and ineffectiveness were expanded to include ‘potentially effective’, ‘mixed’ and ‘inconclusive’. This is in line with a similar analysis into the effectiveness of conflict prevention programmes in C Cramer, J Goodhand and R Morris, Evidence Synthesis: What Interventions Have Been Effective in Preventing or Mitigating Armed Violence in Developing and Middle-Income Countries? (London: DFID, 2016).

\footnote{42. OECD, ‘Evaluation Criteria’.

A concluding paper synthesised the learning from each report in order to answer the question driving this research: ‘what can work and what has not worked in P/CVE?’. This final study includes constructive recommendations for policymakers, donors and civil society organisations operating in the field.

Results and Challenges

To date, the team has collated 463 unique publications, with a current breakdown listed in the tables below:44

Table 5: Type of Publication and Number of Studies

<table>
<thead>
<tr>
<th>Type of Publication</th>
<th>Number of Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Study</td>
<td>153 (33%)</td>
</tr>
<tr>
<td>Programme Evaluation</td>
<td>99 (21%)</td>
</tr>
<tr>
<td>Research Report</td>
<td>93 (20%)</td>
</tr>
<tr>
<td>Conference Report</td>
<td>15 (3%)</td>
</tr>
<tr>
<td>Policy Report</td>
<td>76 (16%)</td>
</tr>
<tr>
<td>Commentary</td>
<td>23 (5%)</td>
</tr>
<tr>
<td>Blog</td>
<td>4 (1%)</td>
</tr>
</tbody>
</table>

Source: Authors’ research.

Table 6: Research Data Type and Number of Studies

<table>
<thead>
<tr>
<th>Research Data Type</th>
<th>Number of Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>190 (41%)</td>
</tr>
<tr>
<td>Secondary</td>
<td>192 (42%)</td>
</tr>
<tr>
<td>Theoretical/Conceptual</td>
<td>81 (17%)</td>
</tr>
</tbody>
</table>

Source: Authors’ research.

Table 7: Research Methods and Number of Studies

<table>
<thead>
<tr>
<th>Research Methods</th>
<th>Number of Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qualitative</td>
<td>285 (62%)</td>
</tr>
<tr>
<td>Quantitative</td>
<td>9 (2%)</td>
</tr>
<tr>
<td>Mixed Methods</td>
<td>79 (17%)</td>
</tr>
<tr>
<td>No Methodology Given (N/A)</td>
<td>90 (20%)</td>
</tr>
</tbody>
</table>

Source: Authors’ research.

44. Please note that this number is likely to increase to over 500 given that further snowballing of data related to several thematic P/CVE intervention areas will still take place.
Table 8: Research Design and Number of Studies

<table>
<thead>
<tr>
<th>Research Design</th>
<th>Number of Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental (Primary)</td>
<td>8 (2%)</td>
</tr>
<tr>
<td>Quasi-Experimental (Primary)</td>
<td>4 (1%)</td>
</tr>
<tr>
<td>Observational (Primary)</td>
<td>157 (34%)</td>
</tr>
<tr>
<td>Systematic Review (Secondary)</td>
<td>6 (1%)</td>
</tr>
<tr>
<td>Other Review (Secondary)</td>
<td>160 (35%)</td>
</tr>
<tr>
<td>No Design Given (N/A)</td>
<td>128 (28%)</td>
</tr>
</tbody>
</table>

Source: Authors’ research.

The design and application of this approach was not without challenges, and the team concedes that despite subjecting its methodology to critical review by P/CVE experts in a consultative workshop convened by RUSI in February 2018, the project may still have been susceptible to some shortfalls and inconsistencies.

The team appreciated the difficulties of sourcing data from the outset but were hopeful that there may be greater stakeholder appetite to share information given repeated calls for greater transparency and exchange from donors and practitioners. Despite formal requests to at least 10 donors, none shared unpublished evaluation material. Acknowledgement and thanks for their valuable contribution go to some civil society organisations and research institutes that did provide access to internal documentation. Nevertheless, the dearth of material was problematic.

Given the lack of available peer-reviewed and public evaluations, grey literature was included to accurately reflect the complexion of the P/CVE evidence base. Integrating ‘non-academic’ material, such as journalistic accounts, policy briefs, presentations, practitioner reports and good practice/toolkit documents, allowed a dynamic assessment of prevention activities and facilitated a more in-depth analysis of what was perceived to have ‘worked’ or ‘not worked’. Crucially, it also enabled the identification and tracking of common assumptions referenced and recycled throughout the literature to understand if there is any empirical evidence to substantiate such claims.

Nevertheless, this approach did present challenges. For example, collating relevant grey literature was difficult due to the sheer scope and diversity of content. It was also widely dispersed, making it hard to capture in a comprehensive and systematic way. While the team tried to mitigate these challenges with hand searches, snowballing and our own expert knowledge of P/CVE information sources, it is possible some valuable content may have been inadvertently omitted.

The reliance on English-language documentation likely distorted the review’s findings, creating a potential bias towards Anglophonic scholarship and expertise largely situated in Western (high-income) countries. Consequently, the study’s geographic coverage may not necessarily reflect the true breadth of the P/CVE space, although it is noted that many authors write in English, and donor- and government-funded publications are frequently translated. This means
that important interventions taking place in non-English-speaking countries have largely been captured. However, reductionism may still have been a problem given the challenges of including innovative or effective activities outside mainstream sources and search engines, especially locally led initiatives at the grassroot level that often receive little external attention and rarely have the capacity or budget to publish or disseminate their monitoring/evaluation outputs.

Relying on institutional and organisational websites also potentially undermined the objectivity of the search and retrieval process by introducing a degree of human bias.\textsuperscript{45} As Richard Mallett and colleagues argue, divergent search functions and the unintentional exclusion of relevant sites means ‘potentially high numbers of pertinent studies can be missed’.\textsuperscript{46} Using the team’s subject-matter expertise, an extensive stakeholder mapping was conducted to mitigate any oversights, but the scope and opacity of the P/CVE space created significant challenges.

Moreover, systematically distinguishing between primary and secondary-level interventions remained difficult, with certain studies requiring ad hoc arbitration by the team to see if it satisfied the inclusion criteria. These issues are clearly demonstrated in the inclusion of education-based interventions: although activities in the education space are rarely targeted at ‘vulnerable’ audiences and often engage all school-aged youth. As such, it could be considered a primary intervention. Yet, education initiatives included in this review described themselves as P/CVE interventions on the basis that the lack of education is a possible structural factor contributing to VE, radicalisation and recruitment. Even if we subsequently assessed that the projects described were primary-level interventions, they were still included on the basis of our inclusion criteria: they described themselves as P/CVE activities. In contrast, broader programmes tackling racism, bullying or civic awareness with no reference to VE or radicalisation were omitted.

Similarly, the team repeatedly cross-checked the scores of each article to limit any variance, but due to the discretionary and subjective nature of the quality scoring process, imperfection and bias were inexorable. While the quality scoring framework was adapted from DFID’s good practice for evidence assessment, there is also an implicit bias towards peer-reviewed academic content. The citation of existing literature, the specification of research methods and the emphasis on independence and empiricism in a given study are important traits and certainly strengthen its authority, but programmatic evaluations, for example, are not necessarily designed for this purpose. The premise of this method may therefore unfairly score papers that do not meet these criteria, enumerating scores that do not necessarily represent their quality or strength.

Finally, the paucity of independent evaluations and peer-reviewed material has challenged the methodological rigour of the analysis. The approach aimed to mitigate some of these problems, but the team acknowledges that conclusions have sometimes failed to be drawn or have been formed on partial data and are therefore liable to be subjective. As such, all findings need to

\begin{footnotesize}
\begin{enumerate}
\item Mallett et al., ‘The Benefits and Challenges of Using Systematic Reviews in International Development’, p. 449.
\item Ibid.
\end{enumerate}
\end{footnotesize}
be viewed cautiously and as an attempt to contribute towards emerging efforts to build the evidence base for research in the field of P/CVE.

Nevertheless, this project provides a valuable resource aimed at strengthening the knowledge base in prevention work, navigating where possible the conceptual, methodological and practical problems prevalent in the P/CVE space, and contributing to improvements in future programming.

*This research methodology has been published in full as Michael Jones and Emily Winterbotham, ‘Research Methodology: The Prevention Project’, RUSI Occasional Papers, May 2020.*

**Michael Jones** is a Research Fellow in the Terrorism and Conflict team at RUSI, examining political violence, conflict economies and the convergence of violent extremism and insurgent militancy in East and sub-Saharan Africa. He has led investigative fieldwork across various countries including Sudan, Kenya and Lebanon, and worked in RUSI’s Nairobi office on a range of projects related to the EU’s STRIVE Horn of Africa and STRIVE II programming.

**Emily Winterbotham** is Director of the Terrorism and Conflict group and a Senior Research Fellow at RUSI focusing on extremism and radicalisation, countering violent extremism and peacebuilding. Between 2009 and 2015, she worked in Afghanistan, most recently as Political Adviser for the European Union Special Representative. Emily has over 10 years’ desk and field experience in an international policymaking environment and is a Deployable Civilian Expert for the UK government’s Stabilisation Unit.