

An imprecise science:

Assessing interventions for the prevention, disengagement and de-radicalisation of left and right-wing extremists

Jacob Davey, Henry Tuck and Amarnath Amarasingam

About this paper

This report examines intervention approaches currently being used to challenge political extremism globally, drawing on 19 interviews with practitioners based in Belgium, Canada, Germany, France, Sweden, the UK and USA. In particular, it looks at the approaches which are being used to de-radicalise and disengage right wing extremists, as well as current strategies being employed to tackle radicalisation and extremism online. It highlights a wide range of different approaches to this work, some key structural challenges facing practitioners seeking to deliver interventions, as well as a number of discrepancies in the scale and sophistication of interventions initiatives globally.

Acknowledgements

This report would not have been possible without the funding support from Jigsaw, research assistance from Charlotte Moeyens, Jakob Guhl and Cecile Guerin, and the time, insights and expertise of our fantastic interviewees.

About authors

Jacob Davey, Senior Research Manager at ISD

Jacob Davey leads on research into far-right extremism and hate crime. His work focusses on mapping the global far-right and the development of new methodologies for tracking far-right extremism. In this role Jacob has advised national and local policy makers on right wing extremism. Jacob's reports include tracking the internationalisation and mainstreaming of right-wing extremism, and online interventions for countering extremism.

Henry Tuck, Head of Policy & Research at ISD

Henry Tuck manages ISD's research programme and leads ISD's policy-focused work to counter online harms in collaboration with a range of key stakeholders, advising leading governments, international organisations and major private sector tech companies. Since starting at ISD in 2013, he has led a series of ISD's education and counter-narrative programmes and conducted research on foreign fighters, deradicalisation and the far-right.

Dr Amarnath Amarasingam, Senior Research Fellow at ISD

Dr Amarnath Amarasingam is the author of *Pain, Pride, and Politics: Social Movement Activism and the Sri Lankan Tamil Diaspora in Canada* (2015). He is the editor of *Sri Lanka: The Struggle for Peace in the Aftermath of War* (2016), *The Stewart/Colbert Effect: Essays on the Real Impacts of Fake News* (2011) and *Religion and the New Atheism: A Critical Appraisal* (2010). Amarasingam is an experienced field researcher, having conducted hundreds of interviews on social movement activism, organisational dynamics, and youth identity in the Sri Lankan Tamil diaspora. He has also conducted interviews with former fighters throughout the former war zones of Sri Lanka in 2012 and 2013 and at the Turkey-Syria border as well as former and current activists of the Syrian revolution. He has also conducted field research in Somalia, Lebanon, Israel/Palestine, and amongst different ethnic and religious groups in North America and Europe.

Copyright © Institute for Strategic Dialogue (2019). Institute for Strategic Dialogue (ISD) is a company limited by guarantee, registered office address PO Box 75769, London, SW1P 9ER. ISD is registered in England with company registration number 06581421 and registered charity number 1141069. All Rights Reserved. Any copying, reproduction or exploitation of the whole or any part of this document or attachments without prior written approval from ISD is prohibited.

Contents

Executive Summary	4
Key Findings	4
Limitations	7
Recommendations	8
Introduction	10
Extremism and Counter-Extremism Interventions	10
Methodology and Sample	10
Findings	11
Offline Interventions	12
Online Interventions	21
Appendix 1: Overview of Intervention Initiatives	30
Appendix 2: Definitions and Typology of Interventions	32
Definition of Intervention	32
Type of Interventions Studied	32
Appendix 3: Methodology	34
Breakdown of Interventions Providers Interviewed	34
Overview of the Interview Process	34
Bibliography	39

Executive Summary

Key findings, limitations and recommendations

In contrast to the growing body of research into the de-radicalisation of Islamist extremists, there is a dearth of intervention programmes and analysis of their effectiveness for the extreme right and extreme left, despite an increasing recognition internationally of the risks posed by these forms of extremism. Despite the emphasis placed on preventing and countering Islamist extremism and terrorism since the turn of the century, other forms of extremism have become increasingly prevalent in recent years.

Incidents such as recent attacks inspired by white supremacist and white nativist ideology in El Paso, Christchurch and Pittsburgh provide clear examples of the dangers of leaving other forms of extremism unaddressed, but also represent only the most prominent cases symptomatic of a much broader and longstanding trend. For example, between 2009 and 2018 in the US, the extreme right accounted for 73% of extremism or terrorism related deaths according to the Anti-Defamation League, compared with 23% for Islamists and just 3% for left-wing extremists.¹

In recent years a number of national governments have taken steps to address the growing threat from the extreme right, from the UK proscription of National Action in 2016 to the September 2019 US Department of Homeland Security strategy that places a greater emphasis on 'domestic extremism' and interventions.² However, despite recent commitments to counter this phenomenon current responses often remain disproportionate to the threat, and there remain crucial knowledge, skill and resourcing gaps around tackling political extremism effectively, particularly on the extreme right.

This report aims to shed further light on the limited number of programmes and initiatives that deliver interventions to prevent radicalisation and de-radicalise political extremists, identify where and why they have been successful, as well as highlight some of the key challenges practitioners continue to face. The project focused on interventions both online and offline, including a range of initiatives targeting individuals across the radicalisation spectrum, with interviews taking place between February and March 2019.³ Although this project sought to study initiatives targeting a range of political extremists, we found that, in line with recent trends, the majority of the organisations (17) solely targeted the extreme right, with one initiative targeting the extreme left, and one initiative targeting political polarisation writ large.⁴ Our findings indicate that, despite the greater emphasis placed on these threats in a number of national contexts, practitioners remain concerned about crucial gaps in the financial and training resources available to intervention programmes, and called for further independent and comparative impact assessments to better understand what is required to build effective responses that are proportionate to the growing threat.

¹ ADL Center on Extremism. 2019. Murder and Extremism in the United States in 2018. Available at <https://www.adl.org/media/12480/download>.

² UK Home Office, 2016, National Action becomes first extreme right-wing group to be banned in UK. Available at <https://www.gov.uk/government/news/national-action-becomes-first-extreme-right-wing-group-to-be-banned-in-uk>

US Department of Homeland Security, 2019, Strategic Framework For Countering Terrorism And Targeted Violence. Available at https://www.dhs.gov/sites/default/files/publications/19_0920_plcy_strategic-framework-countering-terrorism-targeted-violence.pdf

³ For our working definition of extremism please see Appendix 2.

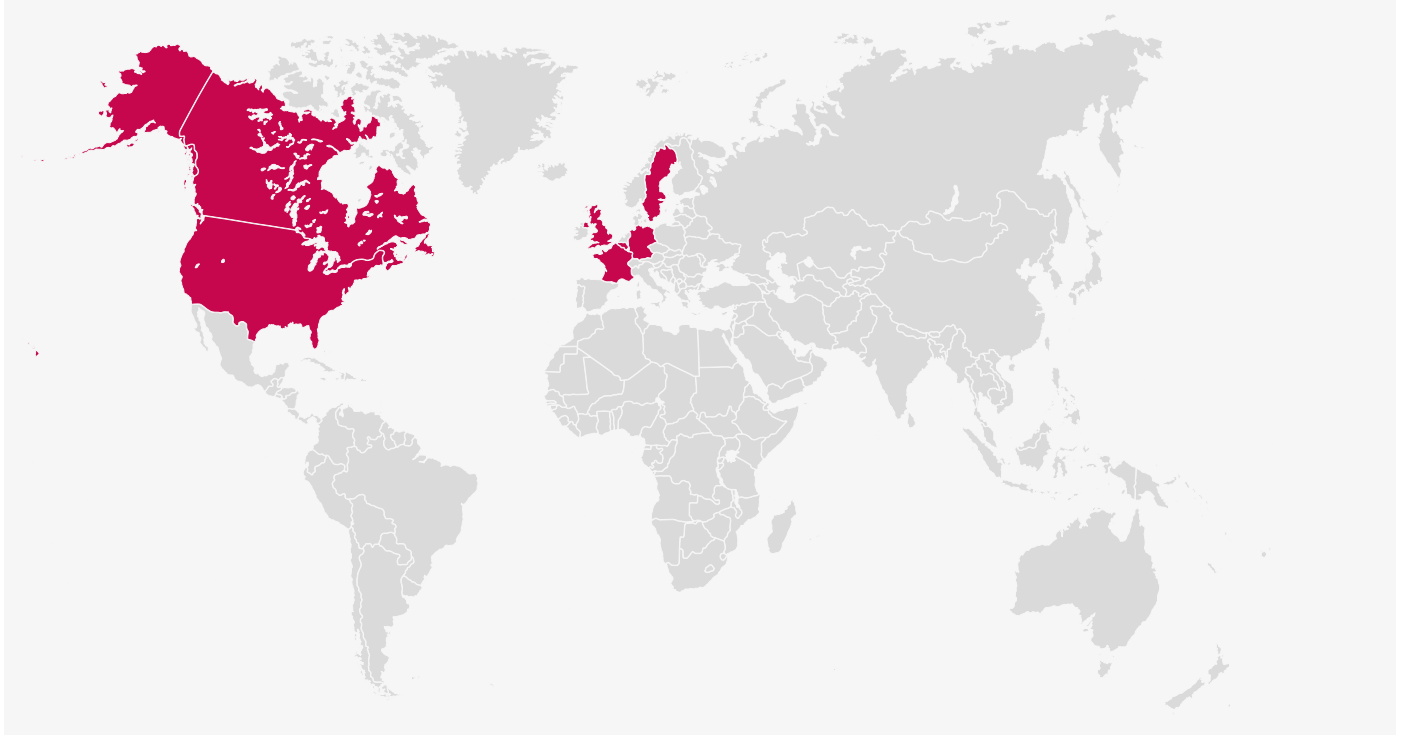
⁴ Our working definition of the extreme right draws on Cas Mudde's definition of right-wing extremism which understands the extreme right to be typically marked by three or more of the following five characteristics: nationalism, racism, xenophobia, anti-democracy and strong-state advocacy. Individuals on the extreme right can be both violent and non-violent, and the ideology has inspired a number of high-profile terrorist attacks. Violent manifestations the extreme right pose a greater cause for concern, however the ideology which underpins such actions is similar regardless of whether adherents commit acts of violence, and accordingly interventions target both the violent and non-violent extreme right.

Our working definition of the extreme left is understood to be left-wing movements which are typically oriented towards extreme political activism involving issues such as animal rights, the environment, pro-immigration and 'anti-fascism'. In particular, these groups are a cause of concern when left-wing ideology is used to justify violence towards political opponents, and hostility towards certain minority communities; in particular anti-Semitism.

Offline interventions often focus on one-to-one mentorship while online interventions include a broader array of approaches, including direct one-to-one outreach, one-to-many outreach, and organised campaigns. This study identified a diverse range of approaches deployed to counter politically motivated extremism, including mentor-based interventions, direct online outreach to radicalised individuals, and semi-automated engagement online. Some intervention providers sought out potential intervention candidates, while others only worked with individuals referred to them from outside services. In some instances, radicalised individuals reached out to intervention providers for support in leaving an extremist group or movement over social media. These interventions often took place across the radicalisation spectrum and depending on the level of engagement with an individual were often highly personalised. Approaches were derived from a number of different fields, from social work to counselling and psychology. This diversity is challenging when considering what constitutes 'best practice' in the intervention space, as it makes the comparison of programmes difficult without a more comprehensive, systematic review of the broad range of initiatives currently operating across these different types of approaches.

1. **Offline practitioners are increasingly bringing their work online because it can help facilitate the identification and initial engagement with radicalised individuals, and lowers the barrier for entry into an intervention.** Practitioners who typically work offline are increasingly, on their own initiative, making use of social media to promote their services and engage with radicalised individuals. This was thought to bring a number of benefits: it allows providers to reach out more easily to radicalised and radicalising individuals; it allows for a degree of distance or anonymity that can reduce the perceived risks for radicalised individuals to engage with intervention providers; and it reduces costs, allowing more interventions to be delivered. However, such online outreach is usually conducted by intervention providers independently, in an ad hoc and un-coordinated fashion, rather than via structured programmes. In the long run, online outreach and more substantive, sustained offline support should be combined to optimise effectiveness.
2. **Online interventions can take a variety of forms, but owing to the nature of unsolicited online engagement may be more effective when the practitioner is engaging with less radicalised individuals in a preventative capacity.** Intervention providers engaged in online initiatives often characterised their work as operating with less radicalised individuals, and felt that online engagement, either through campaigns or direct

Figure 1, Breakdown of where the initiatives included in this report were based



outreach, was often a less appropriate tool than offline engagement for dealing with significantly radicalised individuals who are sceptical and less open to opposing viewpoints. However, this was not true in all cases, and some practitioners, particularly those who were former extremists themselves, did reach out directly to radicalised individuals online as a starting point for a more in-depth, sustained intervention. The appetite for engaging with significantly radicalised individuals typically shifted, according to the practitioner.

3. Technical changes to a platform, or shifts in platform policy, have the potential to disrupt negatively the methods employed by a range of counterspeech practitioners. As with online extremist activity, online intervention methodologies are highly dependent on the infrastructure of a particular technology platform (such as algorithms which prioritise responses in comment threads, or the ability to message non-contacts directly). Accordingly, adjustments made to platform functionality to counter various forms of online harms may have the unintended consequence of frustrating efforts to counter extremism online. As a result, any changes should be thoroughly trialled and tested to ensure net positive effects.

4. There is limited consensus on how to measure success in interventions due to a critical lack of systematic or independent impact evaluations. The process of gauging whether an intervention has been successful is often inconsistent and unsystematised. Many of the intervention providers interviewed, even those working in established programmes, described basing their judgements about effectiveness on 'gut instinct'. Analysis of online campaigns highlighted how existing frameworks for measuring online initiatives are often unable to provide deeper insight into the true effectiveness of an initiative, largely focusing on reach and engagement metrics while failing to measure attitudinal or behavioural shifts. Furthermore, non-governmental initiatives were rarely evaluated independently, resulting in a limited evidence base to demonstrate their impact. Indeed, only one of the non-government-led initiatives studied in this programme has been subject to

independent evaluation. However, some interesting options for evaluation were raised that could be useful to explore further, including the use of machine learning and network mapping to examine an individual's online behaviour before and after an intervention.

5. Intervention providers, and in particular those operating openly online, are not always fully aware of the extent of the potential risks to their own safety and wellbeing, or lack the resources or support required to mitigate these effectively. Engaging in interventions can put practitioners at risk of harm, including online harassment, death threats and even physical assault. Social media has enabled extremists to engage in severe harassment of their opponents, often without consequences. Intervention providers tended to account for this, but preparations for such incidents were inconsistent, and remain ad hoc in more unstructured initiatives. Many intervention providers, particularly those working in a preventative capacity, had not considered the possibility that they might become targets, and accordingly did not take steps to mitigate this risk until after they had become targeted.

6. There are notable discrepancies between the scale, programme maturity and availability of government support between intervention programmes in the US and the other countries studied. Intervention providers operating in the US argued that intervention work is significantly underfunded, limiting their ability to deliver interventions effectively. Many interviewees deliver interventions in their spare time, relying on other forms of income, and expressed a desire to deliver interventions full time while paid a living wage. US initiatives were notably less systematic than those operating in Europe, and lack ties to complementary support infrastructures (e.g. social services, councillors, law enforcement).

7. Intervention providers need further support, including training opportunities, programme infrastructure and access to resources. Many intervention providers interviewed in this work, particularly those operating independently, were

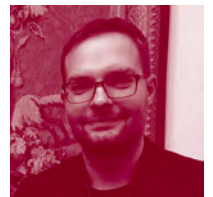
desperate for more access to education, training, best practice and closer ties to related services, and expressed a desire to conduct interventions full time while paid a living wage. Given the diversity and specificity of the types of support required by radicalised individuals, there is scope for expertise to be drawn from a range of fields, including safeguarding, social work, psychology, community-based violence prevention and law enforcement. There was also an explicit request for more support from social media companies, who interviewees felt could be doing more to assist in certain processes, such as evaluating preventative communications campaigns.

Limitations

This research project has demonstrated the wide range of approaches employed online and offline to counter politically inspired extremism. Defining a single 'best practice model' for online or offline interventions is not possible given variations in intervention candidates. A major conclusion from this study is that measuring success (and therefore defining best practice) in intervention programmes is a systemic challenge facing the counter-extremism field.

This study demonstrates that the measurement and evaluation of interventions is too often flawed and unsystematic. Indeed, this is a systemic challenge throughout the counter-extremism field. As Daniel Koehler, an expert in intervention provision, summarised when interviewed for this report:

"We don't even know how to evaluate deradicalisation programs based on the facts... If you look at Germany only, there are about 1,500 CVE [countering violent extremism] programs currently active, I can count on one hand programmes that have ever undergone any form of evaluation, and most of those were not even scientifically rigorous. The level of knowledge about these programs is low."



One reason for this is that incentives may be misaligned for intervention providers to have their work independently evaluated. The counter-extremism sector is highly competitive with limited funding available. Accordingly, publicly discussing the shortcomings of an intervention brings with it the risk that future funding will not be available. Furthermore independent external evaluations are also expensive and time-consuming, and initiatives operating with limited budget and on short time scales might not be able to secure them.⁵

⁵ For suggestions on how this could be achieved see Recommendation 6.

Online interventions rarely measure impact, in part driven by the limited metrics afforded by social media platforms' available analytics, as well as the difficulty of attributing changes to the intervention itself. In many instances, the ambition of a counterspeech or counter-narrative initiative is simply to disseminate a positive message and engage with individuals. However, it is difficult to isolate and demonstrate any long-term impact from this type of approach. Where possible deeper qualitative assessment is necessary. However, the interviews that were conducted for this project suggest that this was conducted intermittently, and without a common evaluation framework, making the comparison of findings between programmes difficult. Online interventions were more measurable when they involved an explicit call to action. This could include the completion of an online course, calling a hotline, or entering into a more structured offline intervention.

Offline intervention initiatives were similarly unstructured in their measurement, often relying more on the gut instinct of expert practitioners than written frameworks. This is in part due to the deeply personal process of radicalisation, and it is possible that over-systematisation could limit the efficacy of an intervention, however there was also a lack of understanding among practitioners where frameworks could be found. A range of these have been compiled, including a recent effort by the RAND Corporation that mapped 35 measurement tools in its Violent Extremism Evaluation Measurement Framework,⁶ however lessons from such frameworks are not always translated into easily accessible resources for practitioners, who often have little formal training in delivering this work.

Recommendations

- 1. Online interventions are cheaper and easier to deliver, and have the potential to reach radicalised individuals at scale, while offline interventions have the greater potential for delivering behavioural and attitudinal change. Accordingly, more interventions should use online outreach and engagement to channel individuals to offline support services.** It became clear through this study that there are opportunities to combine online outreach with offline interventions more effectively. A majority of the offline intervention providers interviewed also conducted online engagements with radicalised individuals, although typically they were unsystematised and conducted ad hoc. Interviewees considered online engagement to be cost-effective and enable greater levels of reach, but warned that if not connected to tangible offline support, they are unlikely to be effective in the long term. Likewise, integrating online outreach can help streamline the provision of interventions to radicalised individuals, and is potentially a useful tool for broadening the reach of intervention programmes. If online and offline initiatives can be combined more effectively, complementing the large-scale reach of online outreach with the deeper behavioural change brought about by offline interventions, then online outreach can play an important role in generating long-lasting change. In particular, building programmes that systematise online intervention provision and connect to existing networks of intervention providers would be beneficial.
- 2. Radicalised individuals actively seek support from intervention providers online. Making intervention services more visible through online promotion could allow more individuals to find off-ramps from extremist movements.** We found that radicalised individuals do seek support for themselves, and often reach out over mainstream social media platforms such as Twitter and Facebook to individual former extremists they have seen giving media interviews or participating in documentary films. With this in mind, it is important to find avenues

⁶ The Violent Extremism Evaluation Measurement (VEEM) Framework, <https://www.rand.org/randeurope/research/projects/violent-extremism-evaluation-measurement-framework-veem.html>.

to amplify the voices of former extremists through initiatives such as the Against Violent Extremism Network.⁷ We also found that former extremists are sought out by individuals from all over the world, not just their own countries. Amplifying and further connecting the global community of former extremists and organisations working in this space to the appropriate audiences would allow individuals attempting to leave extremist groups to identify, or be connected to, appropriate (and ideally local) intervention providers. A key shortcoming in counter-narrative strategies identified in this study is the difficulty in demonstrating behavioural and attitudinal change resulting from exposure to a counter-narrative. If work promoting intervention services can be incorporated into a counter-narrative strategy, this offers an opportunity for a clear impact metric – people reaching out for support from an intervention provider.

3. **Using technology to automate and facilitate the identification and engagement of radicalised individuals online could increase the scale and pace that interventions are delivered.** As more practitioners bring their intervention work online, consideration needs to be given to how this work can be streamlined and systematised. The practitioners we interviewed observed that the use of technology to help identify individuals for intervention, and to manage the intervention process, was a promising development. While a fully automated intervention process brings with it a range of ethical concerns, the tech sector should experiment with technological support services that have proven successful in tackling other issues such as suicide prevention and mental health support, including the use of chatbots and crisis messaging options to initiate conversations. Potential barriers to scaling this work need to be considered further, including how such approaches can be integrated into more structured and/or government-linked programmes, and concerns over privacy, data sharing and surveillance.
4. **Toolkits and resources that provide practical advice to help intervention providers protect themselves from online and offline harassment are essential to prevent intervention providers**

from desisting their activities. Practitioners highlighted concerns around being subjected to direct threats to their safety, and some incidents of harassment and physical assault were referenced in interviews. Moreover, many described being surprised by the scale and nature of the risks involved in this work. Our research suggests that intervention providers would benefit from access to resources that provides insight into how they can keep themselves safe, online and offline. A toolkit that helps to prepare intervention providers to address the potential of harm to themselves and the candidates they work with would be a valuable contribution to the field. Specifically, guidance on how to prevent doxing and advice on when it is appropriate to involve law enforcement could better mitigate the inherent risks of this work.

5. **Further support is required to allow intervention providers operating small-scale initiatives to improve their systems, processes and range of skills. This includes access to training, evaluation tools and partnerships with a diverse group of practitioners, including other intervention providers, law enforcement, psychiatrists and social workers.** While academics and policymakers often push for the professionalisation of the intervention space, some interviewees (especially former extremists) noted that the lack of a bureaucratic structure actually allowed them to be more effective, and to forge strong personal bonds with individuals wanting to leave extremist movements. However, this is a qualitative assessment, and not supported by substantive measurement and evaluation, and intervention providers also noted that they would benefit from more professional training opportunities across areas of relevant practice (e.g. social work, psychology), and access to evaluation tools to substantiate and validate their work. In particular, partnerships with other practitioners such as social workers were cited as useful where they were formalised in more established programmes. Building infrastructure that can connect the disparate groups of intervention providers, who are often working in isolation, with each other, alternative intervention providers, and tools and training are required.

⁷Against Violent Extremism, <http://www.againstviolentextremism.org/>.

Introduction

6. **Further research into the effective evaluation of interventions is vital. Sufficient resources and expertise must be provided to intervention providers to achieve this, and the removal of any disincentives to in-depth evaluation.** This study highlighted significant knowledge gaps in measuring online and offline interventions in the counter-extremism sector effectively, and interviews with practitioners highlighted how current programmes are rarely evaluated independently. As a result, it is possible that existing initiatives may have little or no impact, and some may in fact prove counter-productive. There are challenges in evaluating interventions externally as programmes often rely on intervention providers' in-depth knowledge of particular cases, and a 'gut instinct' that comes from experience. However, taking these factors into account, interviewees nevertheless expressed a desire to create more concrete and accessible tools for evaluation, and a desire for greater support from social media companies in evaluating their initiatives online. Furthermore, practitioners referenced the difficulty in securing long-term funding for projects, so it is challenging to build a strong evidence base around interventions. Accordingly, it is important for funding bodies to recognise the need to support long-term intervention programmes, and incentivise more in-depth and longitudinal evaluation studies. Furthermore, robust evaluation frameworks should be incorporated into new initiatives delivering counter-extremism work, and funders should incentivise pre-existing initiatives to enter into independent evaluations, without punishing their practitioners for partial effectiveness in a sector where demonstrating effectiveness can be challenging.

Findings

Extremism and Counter-Extremism Interventions⁸

In recent years politically inspired extremism has flourished across Europe and North America. In response, it has become increasingly necessary to develop initiatives designed to counter the rise of extremism. These take place across the radicalisation process, ranging from preventative pre-radicalisation interventions with at-risk groups to engagement with convicted terrorists. Measures include programmes designed to halt the spread of extremism, messaging designed to counter extremist narratives, and programmes designed to help people leave extremist groups and ideologies behind, both online and offline.⁹ This project was designed to look at the broad range of different approaches, models and structures underpinning intervention efforts, with the aim of establishing best practice guidelines in counter-extremism interventions, as well as identifying ongoing challenges to implementing effective programmes.

Our working definition of intervention for this project was 'a product or initiative that seeks to prevent or reverse radicalisation through contact (either face to face or through a communications medium) with individuals who have been, or may be at risk of becoming, radicalised'. Within this broad definition initiatives can be categorised further into three key strands of interventions, depending on the area of the radicalisation spectrum they aim to impact:

- preventative initiatives that seek to stop individuals becoming radicalised in the first place
- de-radicalisation initiatives that seek to facilitate an ideological or attitudinal shift in subjects towards more moderate and less violent points of view
- disengagement initiatives that seek to facilitate changes in the behaviour of individuals, primarily disassociation from particular extremist groups or networks.

Methodology and Sample

This research drew on structured interviews with practitioners delivering interventions online and offline, and was designed to identify best practice in the intervention sector and the challenges practitioners face.

We conducted 19 interviews for this project with a range of different initiatives, including mentor-based schemes and traditional 'offline' interventions, counter-narrative campaigns, counterspeech initiatives and direct online interventions.

Our interviewees included representatives from eight online intervention programmes, six former extremists who act as direct intervention providers and five representatives of organisations overseeing the delivery of interventions and training of intervention providers. Initiatives studied covered seven countries: Belgium, Canada, Germany, France, Sweden, the UK and the USA.

For a comprehensive outline of our methodology, definitions and typology of interventions see Appendix 2.

⁸ For a review of de-radicalisation efforts please see Appendix 2: Definitions and Typology of Interventions.

⁹ See Koehler (2017) and Holmer, Bauman and Aryaeinejad (2018).

Offline Interventions

This study draws on two distinct sets of interviews, one with practitioners working on offline programmes and one with those involved in online interventions. We have drawn key findings from each of these in the following sections. Unless stated otherwise all quotes are from interviews conducted for this study.

Offline Interventions¹⁰

Interventions take place across the radicalisation spectrum, from preventative engagement with individuals who are at risk of radicalisation to disengagement work which seeks to help violent extremists exit movements

Interviews revealed there is no single clear definition of 'intervention' among practitioners. Discussions around definitions were often shaped by the overall objectives and end goals of an intervention provider's work, be that impacting disengagement from a movement or maintaining regular contact to build trust and hopefully lead to de-radicalisation. As one practitioner remarked: 'I don't have a singular definition, but I think of it as starting a conversation with people who want help... It's about getting them to open their minds to new things, and then from there you can start further dialogue to build on.'

For other intervention providers, it was also important to consider where on the spectrum of radicalisation and de-radicalisation the individual sits, as the approach must be tailored to suit their current needs. As Shannon Martinez, an intervention provider and former extremist from the US, says: 'For me, the idea of intervention begins with the idea of connecting with people who are still actively engaged in the movement, as opposed to disengagement work, which is where people have already sort of set down the ideology, or are ready to set down the ideology.' Daniel Koehler, a founding director of the German Institute on Radicalisation and De-radicalisation Studies, who works extensively on training intervention providers, supported Martinez's approach. According to Koehler, intervention work is 'addressing actively involved extremists at different stages in the radicalisation process with the goal of helping them

leaving their environment behind and achieving a sustained disengagement or de-radicalisation process'.

This work does not have to be delivered to individuals who are actively radicalised either. Intervention provider and former Ku Klux Klan member T. M. Garret notes that interventions can have a role in preventing radicalisation:

"I have for example one case from a professor who just called me a couple of months ago and said, 'TM, I need your help.'. She said I've got this family, they are friends of mine, and their kid... he's not a Nazi yet – but they felt he was on his way... It's not just getting someone out, but also prevention work."



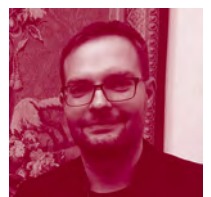
Systematised intervention programmes often bring a range of benefits, including professionalising intervention provision and ensuring intervention providers have access to resources that can support the successful delivery of an intervention

The necessity of systematisation

Interventions are taking place in both systematic and structured programmes, and on an unsystematised individual basis. Intervention providers were asked how they decided what form an intervention would take, including the role of risk assessment tools. This revealed that the process of gauging what support an individual requires is highly personalised, and highlighted the unsystematised nature of a number of programmes.

Some practitioners thought there was a danger associated with the continued informality with which interventions currently take place. Daniel Koehler remarked:

"Intervention is highly specific actually. Most countries start out with a number of professionals, psychologists or social workers, or imams"



¹⁰ This section highlights the findings from interviews with intervention providers and representatives from more established intervention programmes. We held interviews with six intervention providers who are former extremists based in Canada, Sweden, the UK and the USA; and representatives from five established interventions programmes in Belgium, Canada and Germany. The following sections highlight key findings synthesised from these interviews.

or police officers and [believe] that they could process all the information, slowly specialise and then dive into the fieldwork and casework. I found that to be highly risky, indeed intervention work is a specialisation in itself. We need specific BA, MA, PhD-level course[s] to give people a chance to study what researchers have accumulated over decades...We would not do anything else in other fields be it doctors or lawyers. We would make them study, ensure they understand the concepts, and help them move up from less serious cases to more complex ones, let them acquire street-level knowledge. This is very risky work.”

In Koehler’s assessment, the increasing number of cases across a range of extreme ideologies has led many governments and policymakers to rush ahead and pilot programmes without adequately assessing risks, incorporating research and best practice into programming, or considering unintended consequences. This suggests that there is a need for greater systematisation not only for independent intervention providers, but also for government-led programmes. However, questions remain over how this can be realised when there is limited funding available, which was a recurrent complaint among intervention providers in the US.

Martinez argued that former extremists would benefit greatly from this kind of professionalisation. Instead of engaging in intervention work in a haphazard and informal way, many formers would prefer to combine their in-depth understanding of the processes of radicalisation, de-radicalisation and disengagement, and knowledge of ideology, with rigorously tested strategies and techniques informed by research. As she notes, ‘If I had unlimited money, I would love to see programmes to be able to train people to do interventions and disengagement work, particularly formers, because I think formers are particularly suited to have credibility.’

However, this desire for professionalisation is not universal. Some practitioners cautioned against an over-reliance on rigid structures and frameworks

such as risk assessment tools, warning that this might result in individuals being channelled towards support that was inappropriate for them. It is important to differentiate between the use and effectiveness of tools. Rather than the tools themselves being an issue, the interviews suggested that the problem lies in the inconsistent availability and accessibility of these tools for non-experts, and the potential misuse of them by practitioners.

The importance of needs-specific interventions

More established initiatives often take a multi-agency approach whereby practitioners consider a range of different support models for individuals referred to them. These intervention pathways are diverse and can involve therapy and psychosocial support, as well as mentoring. A practitioner working on a government-led initiative noted that in some cases an intervention primarily focuses on connecting at-risk individuals with a stable community group. Indeed, statistics provided by the UK’s Channel programme highlight that 45% of individuals referred are ultimately passed on to other services, including community groups, healthcare, local authorities and educational or employment support, demonstrating how extremism-specific interventions are not always appropriate tools for engaging with people who appear radicalised.¹¹

Most participants recognised the importance of establishing partnerships with a range of social services, observing that the impact of interventions is limited if they take place in isolation. In some cases, co-ordination takes place through informal links with other services. Brad Galloway described how he co-ordinates with a social worker to decide if individuals need an intervention in borderline cases: ‘I will ask a social worker, get them to confidentially look at the message, and if they think the person is reaching out for help, perhaps I will send a message back.’

This highlights how intervention initiatives must be focused on the needs of the individual, and that efforts led solely by one type of intervention provider, such as formers, may be inappropriate for addressing the complex and intersecting needs of a radicalised person.

¹¹ Home Office. 2018. ‘Individuals Referred to and Supported through the Prevent Programme, April 2016 to March 2017.’ Statistical Bulletin 06/18. 27 March. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/694002/individuals-referred-supported-prevent-programme-apr2016-mar2017.pdf.

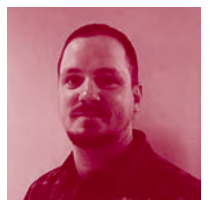
This has implications for intervention providers working independently and without established institutional ties to a range of practitioners. Building infrastructure that facilitates these interdisciplinary connections, even in a semi-formalised fashion, can be an important resource to ensure that radicalised individuals receive an intervention which is appropriate to their needs.

Referrals for interventions come from a range of sources, including at-risk individuals themselves

Depending on how professionalised a programme is, there may be established institutional referral channels. The sources for referrals varied, and included family members, schools and youth groups, among others. Statistics published by the UK's Channel programme listed sources for referral as educational institutions (32%), the police (32%), local authorities (12%), healthcare institutions (8%), the community (4%), family and friends (4%) and prison services (5%).¹²

Radicalised individuals sometimes seek help themselves, suggesting that boosting awareness of available support directly within extremist groups could prove effective, for example with online advertisement of intervention services. As T. M. Garrett and Ivan Humble (an intervention provider and former member of the far-right English Defence League) pointed out, those seeking support may reach out unprompted to intervention providers mentioned in the media or on social media. Brad Galloway noted that intervention providers may receive hate mail for their de-radicalisation work and, on occasion, they decide that these individuals could be worth engaging:

“Someone might reach out from an email that says ‘Blood and Honour 1488’ and tell you ‘fuck you’. Maybe that guy is reaching out, maybe he needs help, but perhaps he’s not in a place where he is ready to talk yet. Oftentimes, people will reach out with a threat, but it’s clear that they want help.”



The nature of the referral process appears to influence a programme's success rate. If a particular intervention provider or intervention programme only deals with individuals who come to them, their success rate will likely be higher. If referrals come from a wider variety of sources, the intake and success rates will likely be lower. This suggests there could be value in finding avenues to signpost the support available from intervention providers to individuals who need support and encourage them to reach out proactively. However, such

¹² Ibid.

an approach comes with the risk that it could open up intervention providers to harassment from committed extremists and trolls, and accordingly would need to be conducted with appropriate support in place, including security and mental health supervision to safeguard practitioners.

Many intervention providers noted that if not carefully managed the referral process could be part of the problem and unintentionally serve to further alienate individuals. Interviewees noted that sometimes parents reach out simply out of fear, or without understanding the full context of what their children are reading or consuming online. Brad Galloway recalls stories of parents reaching out after catching their children looking at websites with far-right symbols, or websites on how to convert to Islam. Shannon Martinez noted that sometimes these referrals arise from employers and others who (understandably) may not have an in-depth understanding of radicalisation. Distinguishing false-positive referrals with legitimate concerns can be difficult, and may require more multi-sectoral approaches to ensure that individuals in genuine need of support are not missed.

Once a referral is received, it is important to gauge when it is appropriate to make contact. Intervention providers take a long-term view, and often reach out to individuals at different times to check in and assess their progress. It is important for intervention providers to understand and recognise when an individual is not ready for an intervention. This insider knowledge – perhaps even an instinct based on their life experience – is a unique value that former extremists can provide. Just because an individual is not currently ready to enter into an intervention, it does not mean that the intervention provider should ignore them completely. Instead, the intervention provider can wait several months and follow up. As Shannon Martinez argues, they may reach out again:

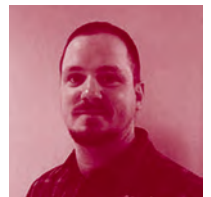
“You know, 6 months from now and say, ‘Hey, hope you’re doing alright, we are still here if you need us.’ I come from the idea that you cannot impose help on people who don’t want it. You are not going to be able to help people



who don’t want help. So, for me there is no harm in starting an intervention where you’re like, ‘You know we’re out here if you need us, contact us.’ If you hear nothing then check in every now and then.”

For more established interventions, referrals can be much more formalised. Often these come from established institutional partnerships with organisations such as prisons, schools and law enforcement agencies. In these instances, the referrals process is often more professionalised and can draw on infrastructure including hotlines and online referral forms. A representative of a government programme in Belgium described this structured process:

“Basically what we have developed is a guiding questionnaire that all the partners have agreed upon to use, so when a phone call or a conversation comes in we all start from the same basis in trying to create an image of what is really happening. From there we bring it to the Partner’s Roundtable which gathers a really broad network of both people looking at it from a professional perspective, but also people knowing people, and being able to add value to the context and how we should read or understand someone’s statements or behaviour.”



There is therefore some commonality between established intervention programmes and smaller scale initiatives, with both taking referrals from institutions and the members of the public concerned about particular individuals.

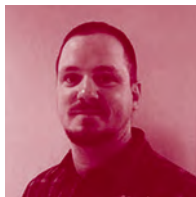
The key variation in interventions appears to be around referral infrastructure and whether intervention providers actively seek out radicalised or radicalising individuals. Although some intervention providers solicit referrals based on their reputation and personal connections, establishing infrastructure that can streamline referrals to these intervention providers could prove successful in ensuring their caseload is managed effectively. Two practitioners from established

intervention initiatives flagged the role of online referral mechanisms as useful, suggesting they could be better harnessed to support individual practitioners.

The lines between online and offline interaction are becoming increasingly blurred, with individuals delivering offline interventions increasingly initiating their interventions online

Offline interventions typically take the form of a mentoring-style relationship between an intervention candidate and their intervention provider, while online interventions tend to focus on opening a dialogue with at-risk individuals and counterspeech. However, the boundaries between these two domains are becoming increasingly blurred, with most of the intervention providers interviewed also taking referrals online as well as initiating contact with apparently radicalised individuals. Many intervention providers maintain social media accounts, conduct media interviews, and participate in documentary films about their lives that are shared online. Thus people from all over the world find ways to get in touch with them for help, and sometimes feel safer to talk about their experiences anonymously through online personas. Brad Galloway observed:

“Some people feel more comfortable in an anonymous context..., the same way that extremist groups often talk to each other, these people trying to exit don’t want to meet in person.”



Online outreach also allows intervention providers to overcome logistical barriers. As Brad Galloway notes:

“Often it is remote, like somebody who’s in Toronto and I’m in Vancouver, the likelihood that you’ll meet in person is pretty low, because we don’t have the resources. I don’t have a guy in Toronto, for every city or community we don’t have somebody who can be there. So, online [communication] is a very good asset... because we can actually talk with people everywhere, including Europe. Community engagement face

to face is super important but we should be using the internet to our advantage.”

This demonstrates that maintaining a social media operation is becoming increasingly important for intervention providers, a number of whom identified the potential of establishing an online presence as a means of encouraging more radicalised individuals to seek support. Mirroring extremist recruitment, the use of social media has a number of benefits for interventions. It allows practitioners to reach out to radicalised and radicalising individuals with ease, and provides a degree of anonymity (often essential in building trust with individuals who are cautious about leaving a movement). It also reduces costs, allowing interventions to be delivered where previously it would have been logistically unfeasible. The maintenance of a public profile also allows an intervention provider’s digital presence to become a beacon attracting those who need help.

This further demonstrates that dividing interventions between online and offline spaces may be artificial. The online outreach identified in this study only took place informally, without a common framework for measuring success in online engagement, and there remains the need to add more rigour to this work, potentially through training for practitioners on safe social media practice. Furthermore, there could be barriers to integrating this work into initiatives that are government-led or supported, such as considerations around online privacy, data handling and surveillance. Accordingly, the merits and challenges of using digital communications to channel individuals towards support services should be explored further.

The length of an intervention depends on the needs of the individual in question, with highly radicalised individuals often requiring sustained support to de-radicalise

Most intervention providers agreed the length of an intervention is highly individualised. Some cases require significant and sustained support over a long period, while others require much more transient attention. According to Daniel Koehler, there is no set standard, but ‘from my own experience, it will take about three times

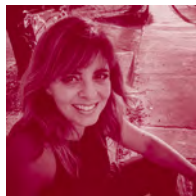
the time a person has spent in a radicalisation process'. Brad Galloway suggests:

"It's a matter for the person who wants to leave. It's about disengagement from the group, people are disengaging but sometimes struggling with ideology problems in their heads, so that can take a while. They reform, go back to their communities, dress normally, go back to work, but they still have problems and they want to talk about them."



Shannon Martinez agrees, commenting on a current case making promising progress:

"It's totally case dependent, and some people need a lot of support. There is a young woman I'm dealing with right now and there are times when what she really needs is some ongoing mental health support, but she doesn't have insurance and she doesn't have access to that, so at this point she and I speak a couple of times per week."



While all of the providers we spoke to recognised that there was no standard timeframe associated with the intervention process, there is a need for more research and tools in this area. Martinez says, 'There is no sense of identified stages and I think that might be helpful, at least a helpful tool.'

Martinez also notes that funding agencies often see trauma and healing as linear but 'it's more likely a spiralling effect'. Individuals can require intervention at numerous points after they have left a movement, depending on the pressures they are currently facing. Accordingly, we should consider intervention candidates as requiring ongoing attention and care.

Former extremists can offer credibility and facilitate the requisite trust building for de-radicalisation

De-radicalisation is a delicate and potentially dangerous process requiring a significant amount of trust between the intervention candidate and the intervention provider. Formers are uniquely positioned to build trust and can relate with intervention candidates through shared experiences and trauma. 'I think that building trust with someone requires you to be vulnerable and be revealing of yourself,' Martinez says, 'so being able to share your struggles, the things that you struggle with, the idea what's happening in your own life is important.' Galloway agrees, 'that's one of the bonuses of having formers involved, it's that I speak the same language as these guys, they will know the jargon', emphasising the importance of having an intervention provider who is credible to the candidate.

The use of formers is not essential; government and non-government initiatives also use a range of alternative providers including social workers and counsellors. Here, the key to trust building revolves around transparency, including the intervention provider's personal experiences, sources of funding, and any 'reporting' requirements this might entail. Most interviewees stressed that intervention providers have an absolute responsibility to disclose their funding sources and any related requirements at the outset in order to prevent undermining trust later on in the process. If this principle is violated, candidates can feel betrayed or manipulated, increasing the chance of recidivism. Martinez notes:

"You are talking about a subset of people who have highly attenuated bullshit meters, they all have street smarts, they have been leading violent lives, they are really good at reading people. The idea of being completely forthcoming with who you are, who you affiliated with, what your responsibilities are in terms of reporting [is important]. To me that first step is like just being completely open and upfront and honest and if you are not able to do that, you should probably not call yourself an interventionist."



Intervention providers emphasised the importance of setting clear boundaries about what individuals can and

cannot share. It demonstrates that there is an ethical, moral and professional component to the intervention process. A factor repeatedly flagged as essential to the trust building process was the adoption of a non-judgemental approach. T. M. Garret summarised:

“Someone who puts a swastika on a school, all these people have valid reasons why they do so. That reason may not be valid to you, or to me, or to most people. For them, they had a very valid reason for their actions. I compare it to the child who sees a monster under their bed. You have a toddler, three or four years old, who calls their mother or daddy and says they’re afraid that there’s a monster under their bed. You know the monster is not real, but for the child the monster is very real. It’s the same compassion we have to give to everyone.”



For more established or institutional initiatives, building trust can be more difficult. Some use formers, such as the Violence Prevention Network in Germany, although they note that they do not actively promote this. Marian Misdrahi (from The Center for the Prevention of Radicalisation to Leading to Violence in Montreal) told us that in many cases building trust begins by having an open discussion in person, where the reason for intervention and the nature of the programme is revealed.

Once trust has developed, the intervention provider is able to start exploring an individual case. This can take a number of different forms. Ivan Humble suggested this could be as simple as ‘sitting down with someone and having a conversation with them’, while others pointed to more specific tactics. T. M. Garret highlighted a process based on ‘motivational interviewing’, which involves focusing on a candidate’s particular motivations for change.¹³ Practitioners in the Violence Prevention Network, meanwhile, prioritise what they characterise

as a ‘pedagogical approach’, which involves training a radicalised individual around psychological approaches, cognitivism and psychology, providing them with the tools they need to engage critically with their own radicalisation.¹⁴ All individual offline practitioners interviewed suggested that direct intervention does not focus on trying to change an individual’s attitudes immediately, but rather on giving them the tools with which they could start to independently question and unpack their ideological beliefs.

Success is difficult to define; its measurement is unsystematised and often highly dependent on the needs of the individual taking part in an intervention

A successful intervention ultimately results in an individual no longer adhering to extremist ideology, or engaging with extremist groups. Additionally, expressing other positive behaviours, such as reconnecting with friends and family, gaining meaningful employment, and positively interacting with groups the individual was previously hostile to, are desirable. However, although the intervention providers interviewed were able to point towards initial indicators of a positive intervention, they do not have a shared definition of success. Brad Galloway observed, ‘It may sound wishy washy, but success is healing.’

There are a number of metrics and factors that could be useful when evaluating the success of a programme, including numbers of individuals reached, levels of recidivism, and numbers of individuals who publicly renounce a movement.¹⁵ However, intervention providers noted that the definition of success depends on the individual circumstances of an intervention. One practitioner said:

“You’d have to assess how deeply the individual is involved in this environment. Is he spending Saturdays and Sundays with a group, or is it seven days a week with a specific organisation? What type of activities [are they engaged in]? You have to assess this, and understand the client.

¹³ Psychology Today. 2019. ‘Motivational Interviewing.’ Available at: <https://www.psychologytoday.com/gb/therapy-types/motivational-interviewing>.

¹⁴ Violence Prevention Network. 2019. ‘Leitbild.’ Available at: <http://violence-prevention-network.de/ueber-uns/leitbild/>.

¹⁵ National Consortium for the Study of Terrorism and Responses to Terrorism. 2016. Surveying CVE Metrics in Prevention, Disengagement and Deradicalization Programs: Report to the Office of University Programs, Science and Technology Directorate, US Department of Homeland Security. Available at: https://www.start.umd.edu/pubs/START_SurveyingCVMetrics_March2016.pdf.

The process is more dynamic than having a tool with boxes.”

This process of gauging whether an intervention has been successful, and whether the individual is at risk of re-offending or re-entering a movement, is often unsystematised. Many intervention providers, even those working within established programmes, relied on instinct to gauge success, rather than using a particular framework or tool.

Despite this apparent lack of structure, it was nevertheless possible to identify a range of metrics to evaluate interventions and factors that increase the likelihood of a successful intervention. Perhaps the most important indicator is the willingness to engage in an intervention in the first place, as this is essential for all other healing processes. Once this initial contact has been made, an important metric that intervention providers referenced was the ability of candidates to be self-critical and recognise the harm that they may have previously committed, as this provides the opportunity for further reconciliation with their pasts, and provides a pathway for future personal development.

One set of success metrics identified by intervention providers relate to a candidate's personal life - whether an individual has reintegrated into society and connected to familial, social or professional support networks. However, when someone has been in a movement for years or has spent time in prison, the goal could be for them to function as independently as possible. In some instances, secondary factors such as finding gainful employment were important metrics of success and precursors to reintegration into society. 'The standard would be for all of them to be fanatical democrats, but it depends on where we start and what base we have', one intervention provider observed.

Another metric of success relates to individuals being able to engage with groups they had previously demonised positively. Intervention provider and former Combat 18 member Nick Cooper referred to a current case:

“The guy I’m dealing with, he’s now ready to go out

and share his story. He’s done some talks with the Muslim community. To me, you haven’t really disengaged if you can’t even go and engage with people. You have to be able to go and talk to people.”



Interventions for countering radicalisation bear similarities to initiatives that help in the rehabilitation of offenders through reconciliation with victims, such as the 'restorative justice' approach.¹⁶

Security and wellbeing of candidates is essential, but smaller initiatives often lack the resources to ensure them

Intervention providers must consider the potential risks of leaving extremist groups faced by individuals, their families or the intervention provider themselves, including violence and threats. In some instances, this may require co-ordination with law enforcement. Risk assessments must be incorporated into every stage of the intervention process, from deciding where to meet to the removal of hateful tattoos, which mark people out as part of a movement potentially making them a target. Secure, private online communication can provide an important alternative, as the individual leaving the movement does not have to worry about risks associated with being seen in public with an intervention provider.

The risk to the individual, especially when leaving an extremist group known for reprisals, could be immense and has to be handled sensitively. Galloway told us:

“I’ll ask them questions about it like, ‘Are you still in contact with these guys? Do you feel like you are at risk? Do you feel like you need protection or help?’... We want to make sure we are providing a safe place for them, so that they don’t feel that things could go wrong in the process, so they actually start sharing.

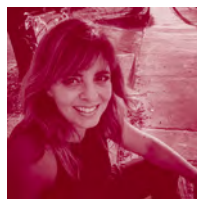


¹⁶ Restorative Justice Council. 2016. 'What Is Restorative Justice?'. Available at: <https://restorativejustice.org.uk/what-restorative-justice>.

Similarly, as they abandon an organisation that provided them with a deep sense of belonging, there may not be another community waiting to embrace individuals who leave extremist groups. Intervention providers can help them cope with this kind of emotional and social limbo, which has the potential to damage the emotional and mental wellbeing of an intervention candidate. One provider commented, 'This [limbo] is a major challenge, as a lot of the people who have been involved in this and leave, they receive a lot of pressure on them from the group they leave, and communities and wider society who don't trust them immediately.'

Other intervention providers take a more detached approach to the issue of security. They may provide knowledge and information about potential risks, but stress that the individual must also take responsibility for their own safety, and learn how to do that in order to remain disengaged in the long term. Martinez argues:

"I do not see that as my responsibility, my responsibility is to try to give them the information that they need to make themselves safe... It would be fantastic if over time there were better working relationships with law enforcement, so you could tell them, 'Hey I've got this person who is disengaged or who is in the process of disengaging, and I'm letting you know not to harass or nothing like that but to just be on the lookout for any weird stuff.'"



Other risks to intervention candidates included psychological trauma and burnout because of the stress of leaving a movement, a factor that can be exacerbated by threats of violence. More established programmes that adopt a multi-agency approach were efficient at addressing this, and with flagging issues to law enforcement, as they tended to have close ties to psychologists and the police. However only some of the independent intervention providers interviewed had similar ties. These were often based on personal connections rather than established institutional partnerships, with the result that there was no standardised level of care and support available to de-radicalising individuals.

US interventions face more funding deficits and lack of institutional support than other countries

Although our geographical sample is by no means comprehensive, our interviews nevertheless pointed to a number of discrepancies between the provision of interventions in the US and the other countries studied. In particular, the availability of funding was a concern. As Shannon Martinez said: 'We have no funding... and we are doing it [delivering interventions], so I'm bartending.'

In addition, the institutionalisation of intervention work differed by location. Intervention providers in the US typically operated independently or as part of small organisations. Although some had ties to law enforcement and additional support networks, this was largely because of personal connections intervention provider made themselves, rather than established institutional arrangements and partnerships.

In comparison, initiatives outside the US tended to be more professional. They were often better funded, although lack of funding was a major concern. The non-US initiatives often had formal ties to a range of partners including social workers, mental health professionals, religious leaders and law enforcement officers. Non-US initiatives also tended to have infrastructure governing the intervention process, as well established channels for referrals of individuals, safeguarding protocols and frameworks for delivering interventions.

In part, this can be attributed to better support from government and a number of other initiatives had close ties to governments through their networks, though only one initiative included in this study was directly government-led. Long-term investment at a national level in local programmes designed to counter political extremism seems to have had a significant effect on the levels of professionalism in intervention provision across Europe.

Online Interventions

This section highlights key findings from interviews with online intervention providers. Three of the initiatives are representatives of the French, German and Swedish franchises of the international counterspeech campaign I Am Here, two are direct online intervention programmes, and the remainder are counter-narrative campaigns, one of which incorporates counterspeech components through encouraging communities to engage directly with people who have been exposed to extremist or hateful content.

A majority of online initiatives focus on preventing radicalisation

All of the online intervention programmes studied arose from a desire to push back against what they perceived to be an increasingly polarised digital landscape allowing hate to flourish. As Mina Dennert, the founder of the movement I Am Here commented: 'There was hatred everywhere, literally, if you had an article about gardening, vacation, every article would get tons of racist comment, illegal stuff, "kill them", "they're cockroaches"... etc.'. The main objective of most of the digital intervention providers interviewed was to offset the hateful conversations occurring online, and shift conversations towards more positive interactions by attempting to counteract echo chambers and reach those not consuming other kinds of messaging. Online interventions took place across the radicalisation spectrum, ranging from the preventative space (Build Up's initiative 'The Commons') to direct engagement with individuals who may be radicalising (such as Project Deconstruct, which intervenes with individuals who are engaging with far-right groups online in Germany).

Seven out of the eight online interventions studied had some sort of preventative component in their work, and it was suggested that online work might be best situated as a preventative measure. Lena Sierts (Project Deconstruct) noted that individuals who had already become fully fledged members of extreme groups are unlikely to benefit from direct engagement online. This was mirrored by one practitioner working on a project using counterspeech and counter-narratives, who intended to reach 'people who, in protest or dissatisfaction or because of disinformation, spread

[far] right-wing populist slogans or vote for right-wing populist parties, but do not yet have a closed far-right extremist world view, as we cannot reach such people anymore'.

This suggests that organisations engaging directly in online interventions focus on disrupting the early stages of radicalisation, while considering de-radicalisation as better suited to offline initiatives. Interestingly, this finding contradicts the interviews with offline intervention providers who stressed the value of online engagement in reaching radicalised individuals. However, this also highlights how different tools for engagement, such as counter-narrative content, have value at different stages of an individual's radicalisation journey.

Several of the initiatives examined sought to promote grassroots counterspeech to extremist conversation online. The three national branches of I Am Here conduct co-ordinated activism, identifying online spaces targeted by extremist groups and promoting positive messaging by taking advantage of Facebook's algorithms to enhance the reach of positive content. Shani Benoualid, who co-created the French iteration of I Am Here, explained: 'The idea is to create a co-ordinated online effort... to spread respect online – regular messages – and to make these messages more seen than hateful messages.' Extremist groups regularly employ co-ordinated activism and platforms' native algorithmic functionality to promote content mirror tactics, suggesting that further analysis of extremist activity could highlight methods that could be appropriated for good.¹⁷

In a similar fashion to I Am Here, other initiatives sought to inspire others to become actively involved in pushing back against extremism and the conditions that give rise to extremism online. Build Up, an initiative that targets individuals exposed to extremism with direct interventions, seeks to attract individuals who have completed an intervention to become intervention providers themselves. One individual working on an online campaign noted: 'We want to motivate and support as many people as possible to become active against right-wing populism.' In these examples, online

¹⁷ Leo Benedictus. 2016. 'Invasion of the Troll Armies: From Russian Trump Supporters to Turkish State Stooges.' Guardian, 6 November. Available at: <https://www.theguardian.com/media/2016/nov/06/troll-armies-social-media-trump-russian>.

interventions go beyond raising awareness around the risks of extremism and tackling extremist narratives, to actively nurturing communities that can provide a civic grassroots response to the spread of extremism online. In some instances, campaigns that primarily have a counter-narrative function can have a secondary objective of boosting community responses to extremism. Empowering grassroots networks also has the potential to backfire. Interventions can unintentionally stoke online conflict and boost polarisation, particularly if groups of online activists decide to engage extremist groups directly. A practitioner working to counter antisemitism in Germany told us:

“We are quite sceptical about counterspeech as a cure... we doubt the impact of it for addressing antisemitism, which is often centred around closed, dogmatic views of the world. It may even backfire or radicalise people further.”

For example, there have been notable instances of left-wing activists actively seeking to counter far-right extremism backfiring and misidentifying individuals online as neo-Nazis.¹⁸ Likewise, misdirected socially progressive online campaigning has been highlighted as a cause of far-right radicalisation.¹⁹

Online and offline interventions use similar approaches, such as employing credible messengers

Many online intervention providers spoke about ‘civil messaging’ and providing people with the material that can help them reach conclusions themselves, rather than dictating how they should or should not think. In this way online and offline interventions are similar. Mina Dennert commented:

“We administrators of the groups don’t tell people what to think or write. We want to encourage people to find information, and help each other with links and facts, to reach their conclusions by



themselves.”

The Build Up initiative actively drew on approaches derived from current best practice in offline mediation, including non-violent communications, empathetic listening, multi-polar discussions, identifying shared values and common ground, and appealing to emotion as opposed to fact. In other initiatives, the use of influencers was seen as an effective and credible way to reach audiences, a trend that matches tactics employed in commercial marketing. A representative from an organisation focused on countering anti-Semitic extremist activity in Germany noted:

“Our best results have used well-known influencers through short, concise and visually appealing but not too complicated contents... [If we had unlimited resources] we would heavily invest in making our contents as visually appealing as possible for each different platform and work with young influencers to actually reach young people.”

This use of influencers demonstrates that the strategy of disseminating online interventions is often comparable to approaches proven effective elsewhere, whether in offline interventions, viral marketing or even extremist propaganda.

Online interventions occur on mainstream platforms, not in the ‘alt-tech’ ecosystem²⁰

All of the programmes surveyed operated on large mainstream social media platforms, including Facebook and Twitter. Some supplemented these with additional tools and software to provide analytics on their campaigns, such as network mapping and project management tools. Approaches were naturally shaped and limited by the functionality of the chosen platforms. Shani Benoualid, who helps lead the French iteration of I Am Here, said: ‘We use Facebook Analytics, but this is an area where progress still needs to be made; you don’t get access to the full range of data that you need.’

¹⁸ Vegas Tenold. 2018. ‘To Doxx a Racist.’ The New Republic. 26 July. Available at: <https://newrepublic.com/article/150159/doxx-racist>.

¹⁹ Nagle (2017).

²⁰ Alt-tech refers to the constellation of alternative social media platforms created and co-opted by extremist communities. For more information see Kevin Roose. 2017. ‘The Alt-Right Created a Parallel Internet. It’s an Unholy Mess.’ The New York Times. 11 December. Available at: <https://www.nytimes.com/2017/12/11/technology/alt-right-internet.html>.

None of the interviewees deliver interventions on the alt-tech platforms (e.g. Gab, 4Chan, 8Chan) that typically attract extremist users. In part, this is explained by the view that online intervention is considered ineffective at engaging significantly radicalised individuals, but it demonstrates a crucial gap in current efforts, and suggests there is a need to pilot engagement models in these broader online ecosystems.

Interviewees commonly identified comment threads on social media as a valuable environment for engagement, as they often attract extremist users. Mina Dennert told us: 'What I saw was that there was a huge spread of racist and sexist material that was spread widely on Facebook... every article would get racist stuff.' This included comment threads beneath extremist and counter-narrative content, with engagement providing an opportunity for activists to discredit and rebuff extremist and hateful messaging, add nuance to particular arguments, and promote positive messaging.

Intervention providers have learned how to use social media algorithms advantageously to promote positive messaging. For example, Shani Benoualid had found that 'Facebook has a hierarchy between messages... when you have a lot of engagement between comments, your comment will arrive at the top'. Online interventions can therefore be highly dependent on the functionality of a particular platform. As a result, technical changes to a platform, or shifts in platform policy, have the potential to affect the efforts of a range of counterspeech practitioners significantly.

Automation can enhance the efficiency of programmes, but should complement and not supersede human engagement

We asked interviewees about their use of automation in interventions, through tools which could either automatically identify individuals expressing support for extremist talking points, or manage the delivery of an intervention. In some previous initiatives, automated tools have been used to identify target audiences for interventions,²¹ but the consensus among practitioners interviewed was that, while automation was useful, it was inappropriate for a machine to be involved in delivering an intervention beyond the initial stages. Of

the initiatives whose practitioners we interviewed for this report, only Build Up's 'The Commons' used automation. Helena Puig Larrauri of Build Up told us:

"Our whole idea is to augment human mediation with automation. One of our hypotheses is that a lot of the offline initiatives for depolarisation are to some degree preaching to the choir... by using automation we are making people aware that they might be part of a polarised narrative... the hand off to the non-automated portion happens when people respond. The monitoring is automated as well."



In this example, automation is used to replace initial human interaction and to conduct often time-consuming processes such as monitoring. However, one interviewee identified the expertise required and the cost of deploying such technology, combined with a lack of understanding of its value at a strategic level, as barriers to further implementation:

"Simply speaking, we genuinely lack financial capabilities to purchase professional software tools, or at least they don't get written into our budgets because the people who make decisions within our organisation or the government bodies that support us lack an understanding for why these tools could help our work."

The use of this technology also had notable limitations. Build Up practitioners deferred from automating any other part of the intervention process, and once individuals had been engaged, the sensitive task of delivering an intervention was handed over to a trained practitioner. Shani Benoualid told us, 'We defend the idea that while some things can be automated, the human factor is really important, so we craft our content without automation.' Primarily this is an ethical consideration, as interventions with at-risk or radicalised individuals can be a sensitive and volatile process, with the potential to impact significantly on an individual's life. If delivered in an unhelpful fashion this could further

²¹ Kevin Munger. 2017. 'Experimentally Reducing Partisan Incivility on Twitter.' 7 September. Available at: <http://kmmunger.github.io/pdfs/jmp.pdf>.

facilitate an individual's radicalisation, or indeed open them up to other problems. However, it should be noted that chatbots are increasingly being employed to address other social harms, such as poor mental health, and as this technology further develops it could have applications in counter-radicalisation.²²

Scale and engagement metrics are not standalone metrics of success

Organisations typically measured the frequency with which counter-narrative material was disseminated, which ranged from twice a day to once every two days in the interventions studied. However, those interviewed cautioned that metrics showing how many people a piece of counter-narrative content has reached are of limited use, and instead sought to identify other measures to determine any impact from their work. A practitioner working to counter antisemitism in Germany observed:

“Reach does not tell you much. Interactions matter, how often contents are shared and with which sentiment. We do look at that very closely. Are the reactions negative or does it foster discussion and further communication? We think it is good when people interact with and comment on our contents.”

Initiatives that focused on conversation-based intervention, through either direct engagement or counterspeech, were notable for their potential scope. These organisations' targets suggested that thousands of interventions could be delivered in a matter of months, partly through automation in the case of Build Up, or because of the significant size of the volunteer bases in the case of counterspeech initiatives. Speaking about I Am Here, Mina Dennert remarked: 'In Sweden [we have] 75,000 members, and in Germany 45,000. France is growing and Italy is growing, and UK is kind of growing. We are 11 groups worldwide.'

The scale of direct online intervention and counterspeech initiatives significantly outstrips the volumes seen in offline interventions. However, the depth of these interventions can be limited, and it

can be difficult to ascertain whether they have had an effect on their target audiences. A campaigner working in a counterspeech initiative commented: 'It's really hard for us to know how involved people are.' In comparison, practitioners working on offline intervention programmes are usually able to judge whether an individual has de-radicalised, even without structured evaluations. Therefore, if online and offline initiatives can be combined more effectively, complementing the large-scale reach of online outreach with the deeper behavioural change brought about by offline interventions, then online outreach can play an important role in generating long-lasting change.

Measuring the impact of online interventions is unsystematic

Participants repeatedly highlighted the ongoing challenges they faced in measuring and evaluating their interventions. Interviewees from organisations delivering counter-narratives highlighted how they capture reach, how long viewers watched a video, and the number of comments on their content. However, when reflecting on this process, participants recognised the limitations of these metrics in evidencing behavioural or attitudinal shifts in their target audiences effectively. Instead, they highlighted the importance of additional qualitative analysis. A German practitioner told us:

“For the numbers, we looked at the business manager to establish metrics such as reach, but qualitatively, we stayed in contact with some people whom we had already consulted for our initial research (from the primary target group) to find out how they had received the contents we published, and what other types of content they would like for the future.”

Active outreach was seen as an important part of the process for nurturing and maintaining supportive individuals who can promote positive messaging through counterspeech, and for engaging individuals who might be radicalising or radicalised. This demonstrates that online interventions are often multi-

²² Kylie Gionet. 2018. 'Meet Tess: The Mental Health Chatbot that Thinks Like a Therapist.' Guardian. 25 April. Available at: <https://www.theguardian.com/society/2018/apr/25/meet-tess-the-mental-health-chatbot-that-thinks-like-a-therapist>; Dillon Browne, Meredith Arthur and Miriam Slozberg. 2018. 'Do Mental Health Chatbots Work?' Healthline. 6 July. Available at: <https://www.healthline.com/health/mental-health/chatbots-reviews>.

modal, combining promotion of a message (the counter-narrative), a civic element (a counterspeech activist community) and an active element (direct engagement with extremists online). However, it is difficult to capture the success of these online programmes and interviewees agreed that outreach cannot easily be measured systematically.

With counterspeech efforts, interviewees observed that it is not always possible to gather the metrics needed to manage a project effectively because of platform limitations. Mina Dennert (I Am Here) noted that with activist networks 'there is really no way for us to know how much they are commenting... it's very hard to know because lots of people participate without telling the group'. In some instances, this is due to the limited access to data afforded by platforms themselves. A practitioner operating on Twitter and Facebook observed,

"On Twitter we can measure our control group, how use of hashtags [by intervention candidates] changed before and after an intervention... On Facebook we can't perform that kind of analysis, so instead what we're hoping to do is to send out a survey."

In an attempt to synthesise learnings more effectively, some respondents had implemented rudimentary evaluation techniques: 'The way we record things is very "artisanal" and not very glamorous... I gathered all the reactions in an Excel sheet.' We suggest that if platforms clearly have a problem with extremist users, and there are activist communities actively trying to solve this issue, then there needs to be much closer collaboration between the platforms and these communities around solving these gaps in evaluation while respecting data privacy. There is also a possibility that gaps in evaluation might be due to a lack of specialist knowledge among campaigners. Where this is the case platforms could work to educate practitioners more effectively in how to evaluate work using the tools available.

Direct online intervention initiatives tended to have more developed measurement and evaluation than counter-narrative campaigns. The Build Up initiative used a campaign management system to measure how many people had been approached for an intervention,

and conversion to a webinar to track the number of individuals who had gone through the programme. The programme is run in one month cycles, which enables evaluation to be iteratively fed into the design of the intervention. The initiative also has an 'ideology' score that considers participants' behaviour on Twitter, and a network mapping system to capture how polarised a participant's attitudes were. This was only possible because of the data access afforded on Twitter, however, and several interviewees noted that there are still significant hurdles in identifying impact metrics on Facebook.

A German practitioner pointed to the use of external evaluation to sense-check their programme when the impact of an intervention relied on a qualitative assessment:

"[We use] external evaluation by two external academics who have much more elaborated knowledge than we do about the current state of research in the field... [We also] network with teachers in schools and universities, because they might be in contact with target group, [and maintain] contact with other groups and institutes... to exchange information about our projects."

This suggests that employing external expert evaluators can be a useful way of adding rigour to an area which otherwise has the potential to be highly subjective.

It is clear from the research that the majority of digital interventions studied had some way of monitoring impact, but the sophistication of these efforts varied. In part, this relates to limitations around functionality, but this area needs further investment and research. The reliance on qualitative methods mirrors the 'gut instinct' approach employed by offline intervention providers, yet while offline intervention providers were (to some degree) happy to proceed in this fashion, most practitioners identified the evaluation of online interventions as a crucial area for improvement.

Engaging in online interventions can be risky for practitioners

Campaigners identified a range of harms including doxing, co-ordinated trolling, and receiving hate mail and death threats. Mina Dennert highlights the potential repercussions:

“We have changed our whole lifestyle, we have changed everything from mailbox to alarm, we got a dog, I open the mail once a week... We have received death threats. Me and my family, and also some of the other administrators, but mostly me. Two people who were in the board before and the admin have been attacked. At one point last year, there were death threats, more than anything people writing fake news about me... I had tons of lies about me and clips on YouTube, just lies, I don’t even look at my Twitter account because it’s all hate there.”



Online interventions are often delivered by small organisations with limited resources, and it might be too expensive for them to implement adequate security measures. However, there are effective measures that people can take to protect themselves online, such as ensuring material with personal information is not readily available. Organisations exist which offer these services to activists,²³ but the findings from the interviews demonstrated that many practitioners are unaware of them. Accordingly, it is suggested that there should be more information available to practitioners about the potential risks of engaging in this work, and toolkits should be promoted which can help them protect themselves cost-effectively.

While the levels of harassment are concerning they are not particularly surprising. Harassment is commonly used by extremist organisations as a means of silencing critics, and social media has enabled extremists to attack their opponents quickly, easily and anonymously. Some practitioners were aware of this risk, but felt they did not have the resources to mitigate the risks effectively. A practitioner working on a combined counterspeech and counter-narrative campaign in Germany observed: ‘We had to realise that doing this type of campaign online is scorched earth... If you look at active counterspeech and support in the comment sections, we were often left with a small group of team members to fight for ourselves’.

What was particularly notable was the fact that several of the interviewees were not initially worried about the risks. As their campaigns grew, they became more aware and started to put measures in place to protect themselves and their staff. This lack of awareness is an issue, as if harassment takes an organisation or individual by surprise then it can be an effective means of forcing practitioners to desist in their activity.

²³ For example: <https://www.securitypositive.com/>

Practitioners perceive the need for significant change at an industry and policy level

When asked to reflect on where there was a need for change, practitioners pointed directly to social media, suggesting that there is a clear need for companies to take more responsibility for the presence of extremism on their platforms. Respondents pointed to the large role social media companies play within society, and suggested that these companies should proactively engage in more work to respond to the challenges. Mina Dennert commented:

“These are illegal materials, there are big groups in Sweden who are allowed to continue, but all they do is post hate speech... What I would really like to see is social media platforms or giants like Facebook, Google and Twitter take their responsibilities.”



In part, this perceived responsibility revolved around questions relating to intermediary liability, and respondents strongly felt that social media companies act as publishers for hateful content. As Mina Dennert said, ‘If I did that as a newspaper publisher I would be sent to jail... They are publishers and should take their responsibilities.’

Respondents said that confronting extremism online was not just the responsibility of tech giants, however, and pointed to the fact that funders and policymakers within government have a limited grasp of the role of social media in society. One practitioner described current challenges with social media in Germany:

“[There is a] limited understanding of social media among policymakers and decision-makers. Sometimes, when I listen to their statements on hate speech and social media I would like people to understand that Facebook and co are a place where opinions are formed and that some of the most successful actors in this space use malign tactics to influence opinion and [manipulate] the facts about a specific situation.”

Practitioners perceived this to be a major issue, as they argued it affects the visibility of their work and the funding available, and that the law is not always effective in addressing these issues. Mina Dennert was critical of the police in Sweden:

“The police need to work better; for example, I don’t know how many reports I made to the police and they don’t even bother to look it up, which is really... I would have to be a lawyer and super involved for them to even look into it.”

The intervention practitioners interviewed did not make suggestions about how these barriers could be addressed. However, our study illustrates the levels of frustration that exist among those working in small organisations over the support they receive from industry and government. Although this in itself is not particularly surprising, it is nevertheless a cause for concern. Without more overt support and encouragement there is a risk that these initiatives may stop their activity altogether. This is also important from a reputational standpoint and suggests companies should consider engaging more with intervention practitioners.

Appendices

Appendix 1:

Overview of Intervention Initiatives

Policies and initiatives aimed at countering extremism and radicalisation often focus on preventing individuals from being drawn into extremist groups, and limiting the impact of those who are already involved through investigative and punitive measures.²⁴ Preventative initiatives can take many forms, including educational and awareness raising campaigns designed to highlight the risks of radicalisation, and community-facing initiatives that focus on tackling societal root-causes of radicalisation. These initiatives are delivered by both governmental and non-governmental bodies, and can be found globally.

In the past two decades a number of highly publicised de-radicalisation and disengagement programmes aimed at violent Islamist extremists have been established in countries such as Egypt, Indonesia and Singapore, albeit with mixed success and differing approaches. In comparison, the first Exit programmes tailored to the far-right began in the late 1990s, initially in Norway, and then in Germany and Sweden.²⁵ Many of the various far-right programmes have been particularly successful and proved sustainable in the long term. Others, despite often achieving promising results, have suffered from a lack of funding and support, and consequently are no longer operating formally.

Considering the potential for violence and other threats from extreme far-right groups, movements and lone actors, Exit programmes specifically designed to disassemble such ideologies are under-represented among de-radicalisation and disengagement initiatives. In comparison, although far-left extremism exists in a number of countries globally, initiatives designed to de-radicalise and disengage far-left extremists are virtually non-existent in Europe and North America, and certainly less common than those targeting the far-right.²⁶

In part, this could be due to the limited threat far-

left groups are deemed to pose in most contexts. As far-left extremists (violent and non-violent) are often deemed to grow in response to far-right groups, rarely seek to impinge on the rights of vulnerable minority communities, and are often non-violent, it can be suggested that they do not represent a priority for governmental or non-governmental intervention.²⁷ There are some intervention initiatives in South America working with the far-left, however the context is very different from that of Europe and North America, so approaches are not necessarily transferable. However, as concerns over far-left extremism continue to grow,²⁸ including around antisemitism and increased violence in response to the far-right, it is possible that initiatives designed to counter far-left extremism and de-radicalise far-left extremists will grow in number.²⁹

Exit programmes attempt to address the underlying extreme ideological belief structures of individuals (de-radicalisation) alongside the practical, behavioural aspects of leaving an extremist group or movement (disengagement) in order to fully reintegrate former extremists back into mainstream society. Such efforts are crucial as fully fledged members of extremist groups present the greatest threat to community cohesion and public safety, and are capable of contributing to the radicalisation of others. This is often achieved through mentorship activities, particularly when targeting youth considered at risk of radicalisation and violent extremism – an approach that shows considerable potential to limit the spread of radicalisation,³⁰ but there is an extremely limited amount of evaluation data available on these initiatives.

Radicalisation to violence can be caused by a variety of individual and societal level risk factors. In his analysis of Somali youth joining al-Shabab, Hassan describes these as 'push' and 'pull' factors.³¹ Push factors are 'the negative social, cultural, and political features of one's environment'.³² These can include poverty, lack of

²⁴ See Veldhuis (2015), Morton and Silber (2018) and Vidino and Hughes (2018).

²⁵ See Koehler (2017).

²⁶ Ministry of Refugee, Immigration and Integration Affairs. 2010. The Challenge of Extremism: Examples of Deradicalisation and Disengagement Programmes in the EU. Available at: file:///C:/Users/SW/Downloads/doc_18843_352415506.pdf.

²⁷ Harry Cheadle. 2018. 'Racist Terrorist Violence Is Not the Same Thing as "Incivility"'. Vice. 30 October. Available at: https://www.vice.com/en_ca/article/qv97v3/racist-terrorist-violence-is-not-the-same-thing-as-incivility.

²⁸ Commission for Countering Extremism. 2018. Annexes to Commission for Countering Extremism Terms of Reference. 20 September. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/742178/Terms_of_Reference_into_Extremism_Study_Annex.pdf.

²⁹ D. Hirsh. 2017. Contemporary Left Anti-Semitism. Abingdon: Routledge.

³⁰ D.L. DuBois and F. Alem. 2017. Mentoring and domestic radicalisation. National Mentoring Resource Center. Available at: http://nationalmentoringresourcecenter.org/images/PDF/Mentoring_and_Domestic_Radicalization.pdf

³¹ As cited in DuBois and Alem (2017).

³² Ibid., p. 4.

education, discrimination, and political and economic marginalisation. By comparison, pull factors are the perceived benefits offered by membership in an extremist organisation, including ideology, friendship and brotherhood, belonging and reputation.³³

The use of former extremists in the intervention process is particularly prevalent. Former extremists have been identified as a potential resource for interventions designed to 'de-radicalise' persons holding extremist beliefs.³⁴ It is posited that former extremists may be essential in these efforts as they have the required credibility to gain the trust of those holding extremist beliefs and to engage them in meaningful dialogue.³⁵ However, de-radicalisation efforts are only one component of a more comprehensive intervention strategy that must prepare individuals for success by equipping them with skills to become productive members of their community.³⁶

³³ Ibid., p. 4.

³⁴ A. Rabasa, S.L. Pettyjohn, J.J. Ghez and C. Boucek. 2010. Deradicalizing Islamist Extremists. RAND Corporation. Available at: <https://apps.dtic.mil/dtic/tr/fulltext/u2/a534160.pdf>

³⁵ Ibid., p. 36.

³⁶ Ibid., p. xvii-xviii.

Appendix 2:

Definitions and Typology of Interventions

Definition of Intervention

The term intervention is often understood by those working within the counter-extremism sector to mean offline mentor-based schemes with radicalised individuals. However, its use is not just confined to this definition. Practitioners working in strategic communications, an approach often deployed to disrupt and counter extremist ideology and messaging, also commonly use the term intervention for communications campaigns.

Accordingly, our working definition of intervention for this project was 'a product or initiative that seeks to prevent or reverse radicalisation through contact (either face to face or through a communications medium) with individuals who have been, or may be at risk of becoming, radicalised'. Ultimately, the selection of a broad definition was validated by interviews conducted during this project, which revealed that among practitioners there is no one definition of what an intervention entails.

Radicalisation to extremist ideologies can take place across a spectrum, ranging from early interest, to committed membership of groups, to the preparation and execution of violent activity. There is no clear pathway through this spectrum, and some individuals can radicalise to the point of violence in a short period of time. This study sought to examine interventions working across the spectrum of radicalisation to capture a range of approaches and lessons learned to understand how these initiatives complement one another and what gaps remain.

It is possible to categorise initiatives further within this broad definition into three key strands of interventions, depending on the area of the radicalisation spectrum they aim to impact:

- *preventative initiatives* that seek to stop individuals becoming radicalised in the first place (such as the No Place for Hate online intervention covered in this report)
- *de-radicalisation initiatives* that seek to facilitate an ideological or attitudinal shift in subjects (such as the work delivered by the Violence Prevention Network in Germany)

- *disengagement initiatives* that seek to facilitate changes in the behaviour of individuals, primarily disassociation from particular extremist groups or networks (such as some of the work delivered by the former extremists interviewed in this programme).

Type of Interventions Studied

We identify three broad approaches or types of intervention programmes, which can operate across the radicalisation spectrum:

- mentor-based schemes and traditional offline interventions
- counter-narrative campaigns and counterspeech
- direct online intervention

Mentor-based schemes and traditional offline interventions

These interventions revolve around direct contact between an intervention provider and radicalised or radicalising individuals with the intention of supporting de-radicalisation and disengagement. Traditionally, this work is seen to take place face to face and offline, although interventions can involve contact between intervention providers and candidates through a variety of mediums, including social media messaging and phone conversations. This approach is often used in government-supported initiatives and established Exit programmes.

Intervention Provider Case Study:

Brad Galloway

Brad Galloway is an intervention provider from Canada who is also a former neo-Nazi skinhead leader. Having been in extreme far-right movements for 13 years, Galloway ultimately left the movement and de-radicalised. He now works as a researcher on youth radicalisation and is directly involved in the delivery of interventions, using online engagement to encourage individuals to leave extremist organisations, and acting as a mentor to support them with disengagement from movements.

Counter-narrative campaigns

Counter-narrative campaigns are initiatives that

disseminate messaging designed to deconstruct, counter and provide credible alternatives to extremist ideology and messaging. This communication is one-way, and consists of the dissemination of communications products such as videos, memes and infographics through social media. Counter-narratives can be designed for audiences at risk of becoming radicalised, or who have already been radicalised, as well as to broader audiences with the aim of building societal resilience to extremism, and can take place online and offline.

Counter-Narrative Case Study: Project Deconstruct

Project Deconstruct is a counter-narrative campaign delivered in Germany that seeks to 'deconstruct' extreme far-right narratives via social media. The project team posts content designed to initiate conversations with individuals interested in or engaging with extreme far-right narratives. The team then provides these individuals with different perspectives and angles on the topic under discussion. If the project reaches an individual deemed too far down the radicalisation spectrum to benefit from online counter-narratives, Project Deconstruct outsources the intervention to an external partner known to deliver tertiary interventions to support disengagement. The project is currently in its pilot phase and will undergo external evaluation at the end of the year.

Counterspeech: Where counter-narrative programmes use content dissemination to dispel extremist narratives and provide positive alternatives to extremism, counterspeech is a two-way process that seeks to counter extremist messaging online through discussion and conversation. This includes structured, semi-structured or informal initiatives, often relying on a volunteer activist base, and aims to influence online conversations, for example, in comment threads, to drown out, dispute and counter extremist discourse.

Counterspeech Case Study: I Am Here

I Am Here is an international counterspeech initiative launched in Sweden by journalist Mina Dennert. The programme counters abuse in discussion threads on

Facebook and has activists join online discussions to defend those receiving abuse from extremist groups, and to push back against hate speech. The initiative consists of a Facebook group that mobilises members to add positive notes on comment sections identified as containing hatred, extremism and disinformation, attempting to rebalance discussions with multiple perspectives.

Direct online intervention

Direct online intervention involves one-to-one outreach over social media to individuals who express signs of having become, or are at risk of being, radicalised, with the aim of bringing about behavioural and attitudinal change. Direct intervention can take place publicly (e.g. through tagging someone in a Tweet) or privately through a direct message. Direct online interventions have elements in common with counterspeech initiatives, with both involving conversation designed to limit radicalisation. The primary difference is that direct online interventions involve engaging in a conversation with an individual identified as being at risk of radicalisation. In contrast, counterspeech focuses on broader conversation-based engagement in public fora with the goal of changing the over-arching narrative of a discussion.

Online Intervention Case Study: Build Up's The Commons

Build Up's intervention The Commons is a pilot designed to address filter bubbles that are destabilising the civic space in the US. This programme operates pre-radicalisation and addresses situations that can give rise to political extremism. The project identifies polarising filter bubbles on Twitter and Facebook, and then uses social media bots to engage with people who are engaging in polarised conversation online. It then deploys a network of trained volunteers to engage these individuals. Once they have been engaged in an online conversation with an intervention provider, they are invited into a call with an intervention provider that highlights the risks of polarisation. At the end of the call, they are invited to take part in an online course that educates them about polarisation. Afterwards individuals are asked if they would like to become intervention providers themselves.

Appendix 3: Methodology

This project sought to combine a mixed methodology incorporating structured interviews with practitioners and desk-based evaluation of other interventions and campaigns. For the interviews, researchers from the Institute for Strategic Dialogue (ISD) scoped 37 individuals and organisations delivering interventions across eight countries. This included individuals acting as mentors for de-radicalisation and disengagement, either working independently, for NGOs and charities, or directly or indirectly for government; representatives from institutions orchestrating the delivery of interventions; and online campaigners delivering counter-narratives, counterspeech and direct online interventions.

Of the 37 initiatives initially scoped, practitioners from 21 agreed to be interviewed, though ultimately only 19 interviews were conducted because the project had a limited duration. Barriers to confirming all interviews included sensitivities around disclosing work that was completed on behalf of government initiatives and the busy schedule of intervention providers.

The individuals interviewed came from seven countries: Belgium, Canada, Germany, France, Sweden, the UK and the US. Although this project sought to study initiatives targeting a range of political extremists, a majority of the organisations (17) solely targeted the extreme right-wing, with one initiative targeting the far-left, and one initiative targeting political polarisation writ large. One of the organisations interviewed was government run. Not all initiatives wanted to be named or directly referenced in this report, and all work has been attributed in accordance to instructions from the interviewees.

Breakdown of Interventions Providers Interviewed

Table 1 (below), provides a breakdown of the types of initiatives interviewed for this project, as well as their location.

While this research pool has some limitations, particularly concerning the evaluation of government-sponsored initiatives, it was nevertheless large enough for us to capture a range of intervention practices and initiate the synthesis of key learnings and approaches employed during interventions. Where possible we have sought to supplement interviews with desk-based research.

Overview of the Interview Process

ISD staff conducted 19 interviews overall with a range of intervention providers and digital campaigners countering political extremism. Interviews were conducted via phone, sound-recorded and later transcribed by ISD staff. Seven interviews were recorded in French or German. Our interviews followed a structured approach, with each interviewer working through a pre-written list of questions. However, interviewers allowed respondents the opportunity to highlight what distinguishes their work from the approaches of other intervention providers or digital campaigners fighting political extremism, and gave them time to ruminate on their practices and convey key messages they felt were not captured during the interview.

The interviews lasted between 40 minutes and 3 hours. The duration largely depended on how precisely our questions applied to the approach to interventions

Table 1 Breakdown of types of initiative included in the project, by country

Type of Initiative / Country	BE	CA	DE	FR	SWE	UK	USA
Government-led offline programme	1						
NGO-led intervention organisation		1	3				
Individual offline intervention provider		1			1	1	3
Online intervention programme			1				1
Online counterspeech programme			1	1	1		
Online counter-narrative programme			2				
Mixed methodology online programme			1				

taken by the interviewee. As our pre-written questions did not always apply to the specific approaches of intervention providers and digital campaigns in their day-to-day work, some questions were unanswered in individual interviews.

Questions and purpose

To account for the different approaches our interviewees used, we developed two questionnaires, one geared towards offline intervention providers (with 32 questions) and one towards individuals and initiatives operating online (with 27 questions).

Questionnaire for offline intervention providers

The questionnaire for the intervention providers was sub-divided into five topics:

1. The Structure and Scope of Intervention Work:

In the opening section, we asked the intervention providers how they got involved in this type of work and how they would define intervention work. We also attempted to get a sense of the context they are operating in. This included asking what their target group is, whether there are people they would not engage with (out of legal, ethical or security considerations), and which partners or organisations they co-operate with during their interventions. We asked them if they trained other providers, and whether they relied on an assessment framework to determine a candidate's suitability for an intervention. These questions were intended to provide us with a general overview of the interviewee's work, and their specific focus.

2. Referral and Intervention Processes: Our second set of questions was designed to provide insights into the referral and intervention stages of each programme, and touched on the measurement and evaluation of interventions. Therefore, we asked our interviewees how and from whom they received referrals, how and by whom it was determined whether an intervention was necessary, which risks could arise, and whether individuals or their guardians needed to consent to participating in the intervention process.

We were interested to learn more about how the initial contact between intervention candidates and intervention providers is established when an intervention process takes place, how long intervention processes usually take, how and how frequently a candidate's progress is evaluated, which intervention pathways are used for risk mitigation, and who decides which pathway is pursued. Finally, we asked how the intervention providers and their organisations measured the impact of their efforts, and what they considered to be a successful intervention.

3. Ethics, Concerns and Risk Management:

The answers to the questions we asked provided insight into the types of risks each programme considers, how intervention providers attempt to mitigate them, and how they manage risks if they are experienced. Potential risks include accusations of stigmatising communities, safeguarding individual candidates and intervention providers themselves, and the danger that the intervention would prove counter-productive and further radicalise an individual.

4. Communications around the Intervention

Programme: Our questionnaire included questions about the communications and public relations aspects of the programmes, in particular towards intervention candidates, their family and their communities.

5. Addressing the Gaps, Long-Term Developments and Aspirations for the Programme:

In our final set of questions we gave interviewees a chance to share their aspirations and plans for their intervention programme, to reflect on any gaps they have identified in their approach, and tell us how they are planning to mitigate these in future. Through our questions we sought to establish the sustainability of intervention programmes, and how best to make use of lessons learned and insights gained.

Questions asked to intervention providers

Topic 1 – Structure and Scope of Interventions Work

1. How did you get involved in this type of work?
2. Do you have a definition for an intervention?
3. What is your/do you have a target population?
 - a. Age;
 - b. Ideology;
 - c. Location;
 - d. None/ Other (e.g. specific institutions).
4. Are there any individuals, cases or issues you are not allowed/ will refuse to engage with (e.g. individuals with a criminal history, children)?
5. Do you have partnerships (either formal or informal) in place with other organisations to help support your interventions work?
 - a. If so, who?
 - b. Please describe the nature of these partnerships (e.g. regularity of meetings, structure of partnerships).
 - c. How do you choose your partners?
6. Does this programme involve the training of intervention providers and/or frontline practitioners?
 - a. How do you assess if an individual is equipped and/ or ready to deliver interventions? Do you employ an assessment framework during the recruitment process and as they deliver interventions?

Topic 2 – Referral and Intervention Processes

7. From whom do you receive referrals? Do you receive referrals directly from [yes or no only];
 - a. The public (e.g. via a hotline); (The individual's family / friends, or the individual themselves?)
 - b. Local law enforcement;
 - c. National Government/agencies;
 - d. Local authority services;
 - e. Institutions (e.g. universities, schools, faith, healthcare);
 - f. Other [ask interviewee to elaborate].
8. How are these interventions received (e.g. medium)?
9. Who decides whether a referral qualifies for an intervention?
 - a. Do you have any partner agencies that contribute to this decision?
 - b. What, if anything do you do with false referrals (e.g. if referrals are being made maliciously, or misguidedly)?

10. When examining a referral do you employ a risk assessment framework?
 - a. Is the same framework employed for every case, or does it differ depending on the specific origins and details of the referral?
 - b. If the individual does not qualify for an intervention, what do you do with his/her referral and information? Is there any follow-up or reassessment?
11. Does the individual or his/her guardian have to consent to the intervention process or is the support able to be delivered without consent?
 - a. No [skip to question 13];
 - b. Yes – how? [proceed to question 12];
 - c. Other (e.g. case-specific).
12. If delivery of interventions requires the consent of the intervention candidate, what do you do if an individual refuses to participate?
13. How does the intervention process start (e.g. forms of outreach)?
14. How long does the intervention process generally last?
15. At what interim periods during and post-intervention do you reassess the individual (e.g. 3, 6, 9, 12, 18, 24 month intervals...)?
16. What types of intervention options/pathways do you use for your risk-mitigation steps?
 - a. Formers;
 - b. Mentors;
 - c. Case-specific intervention providers;
 - d. Psycho-social workers;
 - e. Teachers/ education officials;
 - f. Wrap around support services: e.g. skills development/back to work schemes/ housing, education, etc.;
 - g. Other [ask interviewee to elaborate].
17. How is the type of intervention pathway decided upon (e.g. lead agency, consensus of a multi-agency panel etc.)?
18. What role, if any, do families play in the interventions process?
19. What are the risk assessment factors and/or tools you use during delivery of the intervention, for example, to monitor the intervention candidate's progress?
20. Finally, how do you measure the impact of your programme?

- a. What metrics do you collect to track progress?
- b. How do you define success/ end an intervention?
- c. Do you follow up post intervention?

Topic 3 – Ethics, Concerns and Risk Management

- 21. How do you mitigate risks around accusations of stigmatisation and/or government-led surveillance and censorship? *[This could touch on how the programme is framed, for example, and to what extent any associations with the government are stated or unstated.]*
- 22. How do you develop and maintain a level of trust not only with the individuals who receive interventions, but also with the community at large, as well as with other organisations working in this space?
- 23. How do you ensure an individual is safeguarded during the intervention? *[For example, from communal backlash, burnout, and depending on type and level of radicalization, safeguarded from recruiters and members of extremist movements.]*
- 24. What steps are in place to ensure the intervention provider's security and mental wellbeing is safeguarded? [Note: only applicable if intervention providers are used].
 - a. Do intervention providers generally use their real names or pseudonyms?
 - b. If they use their real names, how is their safety and the safety of their associates and family members secured?
 - c. How is an intervention provider's caseload managed?
- 25. Hands-on intervention work may require intervention providers to be prepared to deal with a range of emotional, mental and other concerns and traumas experienced by the intervention candidate. How do you mitigate risks around the intervention provider being equipped or prepared to do so?
- 26. Interventions may carry the risk of having the opposite effect of what they intend. How do you mitigate chances of an intervention making a situation worse, for example, by further radicalizing an individual?

Topic 4 – Communications around the Intervention Programme

- 27. How do you advertise and communicate the programme to solicit referrals (e.g. types of literature,

website, hotlines facilitation of training and community/sector engagement)?

- 28. How do you describe the programme to the intervention candidate, to their family, as well as to the communities with which he/she is associated?

Topic 5 – Addressing the Gaps, Long-Term Developments and Aspirations for the Programme

- a. Have you identified any existing gaps in the current provision (e.g. skills-training, ideology changes, wrap-around services), or are there any changes you'd like to make if time or money were not an issue?
- 29. Was this programme developed in accordance with other models and/or the learnings and recommendations of other intervention initiatives across the world?
- 30. Are there any other intervention providers or programs which you would recommend?
- 31. Finally, is there anything you would like to add to your interview, in regards to your personal experiences, or any other recommendations you may have for organisations developing interventions programmes in this space?

Questionnaire for digital campaigners

The questionnaire for digital campaigners contained questions relating to objectives, target audiences, software used for campaigns, monitoring and evaluation, processes for organisational learning, risk assessment and perceived gaps in the field of online campaigning against extremism. The questions were designed to investigate how organisations conducting campaigns determine the scope and aim of their work, and how sophisticated the processes they have in place are to ensure campaigns are well managed and continuously improved. We also wanted to get an idea of which aspects of their work campaigners viewed as effective and which were not, and what practitioners in the field of digital communication believe holds them back from making their campaigns more impactful.

To complement the insights gained from the interviews, we conducted desk-based analysis of three digital counter-narrative campaigns, and drew on insights gleaned from intervention initiatives previously delivered by ISD and others.

Questions asked to digital campaigners

1. What is the objective of your intervention / campaign?
 2. Do you focus on a particular phase of radicalisation etc.?
 3. What interventions/ tactics do you utilise to achieve this objective?
 4. What tech/ software do you use to support your work?
 5. How much of your approach is automated/ human driven? Who are the people doing this work?
 6. What is the scale and rate of your intervention/s?
 7. At what stage of an intervention do you set-out the goal and objectives?
 8. What are the objectives and metrics of your campaign/ intervention?
 9. Do you align objectives with specific tactics and performance indicators?
 10. Do you have pre-designed frameworks or templates you use when planning how you will measure the progress of an intervention?
 11. Do you involve the target audience in the design of the intervention?
 12. Do you monitor the progress of the intervention on an ongoing basis? If so, can you illustrate how?
 13. Have you altered or changed the content of an intervention during its delivery period based on monitoring you have conducted?
 14. Do you use any tools or software to support monitoring, measurement or evaluation? i.e., Google Analytics or social listening tools.
 15. Do you involve the target audience during the creation of messaging or content for an intervention?
 16. Do you involve the target audience in the post-intervention evaluation?
 17. Do you have frameworks or templates for recording and analysing intervention impact?
 18. Do you have an organizational process for integrating lessons learned into new interventions?
 19. Do you have a process for sharing organisational knowledge of best practices for digital interventions?
 20. How do you segment the audience of an intervention/ campaign?
 21. Do you employ an independent evaluator to measure the success or impact of your intervention?
 22. Do you carry out a risk assessment of the interventions you deliver?
 23. Do you have structured frameworks or templates for evaluating intervention risk?
 24. Do you use the results from your interventions/ campaigns to inform future work?
 25. Are there any other organisations delivering online or offline interventions who you would recommend?
 26. Are there any gaps you can identify in current efforts to counter political extremism online?
 27. What would you design with unlimited resources?
-

Bibliography

Bartlett, Jamie and Jonathan Birdwell. 2013.
'Cumulative Radicalisation Between the Far-Right and Islamist Groups in the UK: A Review of Evidence'
London: Demos.
Available: [here](#)

Bray, Mark. 2017.
Antifa: The Anti-Fascist Handbook
New York: Melville House Publishing.

DuBois, David L. and Alem, Fasika. 2017.
Mentoring and domestic radicalisation
National Mentoring Resource Center.
Available: [here](#)

Ebner, Julia. 2017.
The Rage: The Vicious Cycle of Islamist and Far-Right Extremism London: IB Taurus.

Holmer, Georgia, Peter Bauman
and Kateira Aryaeinejad. 2018.
Measuring Up: Evaluation the Impact of P/CVE Programs
United States Institute of Peace.
Available [here](#)

Koehler, Daniel. 2017.
Understanding Deradicalization: Methods, Tools, and Programs for Countering Violent Extremism
London: Routledge.

Morton, Jesse and Mitch Silber. 2018.
'When Terrorists Come Home: The Need for Rehabilitating and Reintegrating America's Convicted Jihadists'
Counter Extremism Project.
Available: [here](#)

Mudde, Cas. 2007.
Populist Radical Right Parties in Europe
Cambridge: Cambridge University Press.

Mudde, Cas. 2017.
The Far-Right in America
Oxford: Routledge.

Nagle, Angela. 2017. Kill All Normies: Online Culture Wars from 4Chan and Tumblr to Trump and the Alt-Right
Winchester: Zed Books.

Phillips, Whitney. 2016.
This is Why We Can't Have Nice Things: Mapping the Relationship between Online Trolling and Mainstream Culture
Boston: MIT Press.

Rabasa, Angel. Pettyjohn, Stacie, L. Ghez, Jeremy J. and Boucek, Christopher. (2010).
Deradicalizing Islamist Extremists
RAND Corporation.
Available: [here](#)

Veldhuis, Tinka M. 2015.
'Reintegrating Violent Extremist Offenders: Policy Questions and Lessons Learned'
Occasional Paper. George Washington University Program on Extremism.
Available: [here](#)

Vidino, Lorenzo and Seamus Hughes. 2018.
'America's Terrorism Problem Doesn't End with Prison – It Might Just Begin There.'
Lawfare. 17 June.
Available: [here](#)

Wendling, Mike. 2018.
Alt-Right: From 4Chan to the White House
London: Pluto Books.



Powering solutions
to extremism
and polarisation

PO Box 75769 | London | SW1P 9ER | UK
www.isdglobal.org