UNLEASHING THE POTENTIAL OF SHORT-FORM VIDEO

Strategic Communications for Countering Extremism in the Digital Age

Joe Whittaker[1], Farangiz Atamuradova[2], Kamil Yilmaz[3], Simon Copeland[4], Lilah El Sayed[5], Jon Deedman[6]

Swansea University Prifysgol Abertawe

RUSI

Hedayah Countering Extremism & Violent Extremism

1. Introduction

Recent years have witnessed a surge in the popularity of short-form videos across social media platforms. These concise and captivating videos, typically lasting around 30 seconds, have become an integral part of the contemporary online experience, conveying messages that are "short and sweet." TikTok, a platform solely dedicated to short-form video, has taken the digital world by storm since its launch in 2016. The magnitude of its success is staggering, with over 2 billion total downloads in 2020 (Carmen, 2020) and a remarkable milestone of 1 billion monthly users achieved in 2021 (TikTok, 2021). Recognizing the immense potential of short-form videos, Instagram followed suit in 2020 by introducing "Reels", enabling users to create and share 15-second multi-clip videos on their feeds (Instagram, 2020). Facebook (now Meta), the parent company of Instagram, also embraced this trend, offering users the ability to explore "Reels and Short Videos" with substantial overlap from Instagram and TikTok. Moreover, in a global rollout in 2021, YouTube unleashed "Shorts", a comparable feature empowering users to shoot, share, and binge-watch short videos of up to 60 seconds. Since its debut, YouTube has streamed a mind-boggling 5 trillion Shorts, a testament to the widespread appeal of this format (Wojcicki, 2022).

Strategic communications have emerged as one of the most popular methods of countering the spread of extremism. In 2017, the United Nations Security Council called on member states to develop counter-narratives, highlighting the importance of such messages being delivered and facilitated by range of stakeholders such as community leaders, civil society, and the private sector (UNSC, 2017). For their part, social media platforms have set up initiatives which seek to help the spread of prosocial messages – something not limited to just those countering extremism. YouTube's "Creators for Change" program seeks to "amplify…voices to counter hate and promote tolerance." In 2018, they provided their creators with $5 million for "production and marketing support" (Downs, 2018). TikTok offers the "Creators Forward" initiative in partnership with the United Nations Development Programme whereby young influencers undergo a six-week bootcamp to develop engaging and positive narratives on the platform (UNDP, 2022). Additionally, the "TikTok for Good" program's stated aim is "to inspire and encourage a new generation to have a positive impact on the planet and those around them", which includes assistance with account management, advanced analytics, and algorithmic promotion of content (TikTok, nd).

We can see social media users begin to use this medium for the purposes of spreading prosocial messages to large audiences. Recent examples include a range of TikTok users jumping on the #StopAsianHate hashtag to create narratives discussing and countering racial hatred (Lee and Lee, 2023), as well as #JewToks, a group of Jewish American TikTokers who challenge antisemitism on the platform (Divon & Ebbrecht-Hartmann, 2022). These users are able to leverage an organically built audience with an authentic message to attempt to achieve positive social goals. The rationale of this project is simple: We seek to offer short-form video creators (and those that

1 Senior Lecturer, Swansea University
2 Program Officer, Hedayah
3 Postdoctoral Research Officer, Swansea University
4 Research Fellow, Royal United Services Institute
5 Program Associate, Hedayah
6 Postgraduate, Richmond the American International University in London

are facilitating them) with a document which synthesises key lessons that have been learned in the field of Preventing and Countering Extremism and Violent Extremism (P/CVE)[7] strategic communications. This is a non-technical document titled *Unleashing the Potential of Short-Form Video: A Guide for Creators Making Content to Counter Extremism*. To accompany that document, this report outlines the strategic logic of our approach. Our aim throughout both documents is not to be prescriptive towards content creators. The reason why these creators can be successful at persuasion is because they have built an organic audience through authentic messaging; we do not seek to interfere with this. Rather, these documents present several considerations that creators may find useful to keep in mind when crafting their message.

The report begins by outlining some of the broad knowledge around the idea of mass persuasion, before focusing specifically on lessons that have been learned in the field of P/CVE. This is followed by a synthesis of existing "How To" guides for the creation of strategic communications from a range of policy and practitioner stakeholders. Then, we discuss specific knowledge of audiovisual content, particularly considerations for short-form video content. The report concludes by outlining how stakeholders, including social media platforms, can monitor, measure, and evaluate the impact of this type of content.

2. Strategic Communications and the Difficulty of Persuasion

The history of research into strategic communications, which we define as the use of messaging and other forms of communication to achieve specific goals or objectives, tells us that persuasion is difficult. It is often – implicitly and explicitly – assumed that well-crafted messages have the power to change minds for the better or worse. This is referred to as the 'hypodermic needle model': a simplistic understanding of communications in which audiences were seen as passive receivers of information who could be swayed by powerful media that inject their ideas directly into the audience (McQuail, 1987). This understanding was particularly common around a century ago, particularly in the context of the First and Second World Wars, in which nation-states leveraged propaganda as part of warfare (Ruddock, 2000). However, research has repeatedly demonstrated that the picture is substantially more complicated than this. In his analysis of propaganda from the Holocaust and the Rwandan genocide, Kiper (2015) demonstrates that common discourse often suggests a causal link between propaganda and committing violence, and, in its simplicity, overlooks the complex social dynamics within each community. Rather than changing peoples' minds and motivating them to violence, he argues that propaganda builds on existing cultural factors such as hierarchies, social ideologies, and peers, giving individuals the "green light" to commit violence if they already believed it was the right course of action. This finding is supported by Adena et al. (2015), who analysed antisemitic radio broadcasts under Nazi rule, finding that the effectiveness of the messaging was dependent on listeners' predispositions (i.e., it was effective in areas that were already antisemitic and backfired in areas where antisemitism was low).

7 Preventing and Countering Extremism and Violent Extremism (P/CVE) is a term that carries substantial definitional ambiguity. For the purposes of this report, we define it as attempting to halt the spread of extremism and radicalisation (either by preventing it before it takes hold or by challenging it after it has). Importantly, it consists of non-coercive measures (i.e., it does not include criminal justice or military interventions), but rather, interventions such as counter and alternative narratives; community resilience-building; and digital literacy.

Ellul (2006) also focuses on the role of propaganda, arguing that effective messaging does not directly contradict an audience's existing opinion, but rather builds on their beliefs, stereotypes, and tendencies, to provoke actions. Rather than manipulating a choice within the audience, it is about loosening reflexes. Similarly, O'Shaughnessy (2004) argues that values are impossible to alter overnight. Rather, propaganda seeks to conscript existing values to move towards different judgements and actions. In essence, such content needs to preach to the already converted to stimulate individuals toward their desired action. Fundamentally, it is inherently difficult to persuade any large group of people. Mercier (2020) highlights that authoritarian propaganda, political campaigning, and advertising all show limited effects in changing the beliefs or behaviours of their audience.

Empirical research also bears out the difficulty in changing attitudes. In a neuroscience experiment, Kaplan, Gimbel, and Harris (2016) examined individuals with deeply held political beliefs, finding that when these beliefs were challenged, there was more activation in the parts of the brain that correspond with self-identity and negative emotions. This is an important point to consider in the context of prosocial messages; there is a difference between attempting to persuade someone of something that is relatively low stakes (for example, whether one food item is superior to another) and challenging one's "sacred values", such as deeply held political beliefs or sympathy with one side in an armed a conflict (Ginges et al. 2007; Sachdeva and Medin 2009). When their sacred values are challenged, people are more likely to resist compromise, social influence, and exit strategies (Hamid, 2020).

So far, we have only considered the difficulty in persuading others to change their attitudes. Going a step further, it is equally, if not more, difficult to persuade people to change their behaviour. In a meta-analysis of over 3 million behavioural interventions to promote household action on climate change, Nisa, Belanger, Schumpe, and Faller (2019) found that such interventions have a very small effect at the point of delivery, and no evidence exists that this influence continues afterwards. The most common strategy was an information-based intervention, but these only had very limited evidence of impact.[8]

To summarise, it is clear that it is possible to change minds; people frequently develop new ideas and opinions, as well as engage in new behaviours. However, this section has demonstrated that inducing such change is not an easy or simple process. Rather, persuasion typically happens when strategic communications push at an already open door. Moreover, not all beliefs are created equal; many individuals will develop cognitive protection for their "sacred values", making persuasion around such ideas difficult.

8 It is worth noting that there is some debate on how the findings should be interpreted depending on the statistical methods used. Van der Linden and Goldberg (2020) argue that the effect sizes are somewhat higher, but Nisa et al. (2020) argue in response that their findings of the limited effects still hold.

3. Strategic Communications in P/CVE

Strategic communications often play a primary role in P/CVE. They are operationalised in various ways, which include developing counter- or alternative narratives (Zeiger, 2016; Russell, 2018), using fictional utopian narratives and storytelling (Schegel, 2021), promoting online voices (Helmus, York, and Chalk, 2022), focusing on social ecology (e.g., networks and environments) (Hamid, 2022), and using inoculation to confer resistance to extremist propaganda (Braddock, 2022).

Outlining best practices in counter-narratives in Southeast Asia based on an expert workshop, Zeiger (2016) defines counter-narratives generally as "responses to terrorist claims and narratives", and categorises them five ways: i) peace narratives & inter-faith, inter-ethnic narratives, which encourage general support for peace and non-violence as well as tolerance between ethnicities and religions; ii) religious and ideological narratives, which utilise religious texts and leaders to refute the religious claims of violent extremists; iii) religion as peaceful and non-violent, which reinforces the positive and peaceful nature and teachings of Islam that provide an alternative to extremist narratives; iv) factual counter-narratives that underscore the erroneous aspects of the extremist narratives; and v) emotional and psychological counter-narratives, which address the emotional or psychological incentives (i.e., pull factors) that an individual may feel are a benefit of joining an extremist organisation. In addition, she acknowledges several suggestions by experts that counter-narrative development should recognize challenges in understanding different forms of religious interpretation; be multi-sectoral, multi-dimensional, and disseminated across multiple platforms; consider providing platforms for open and frank discussions; focus on family relationships; and identify credible messengers (ibid, pp. 4-10). In a similar vein, Russell (2018) makes the case for 'localising' communication efforts in P/CVE campaigns. He defends the idea of valuing the local while warning that the gap between academic and practitioners' approaches to localised efforts in developing counter- and alternative narratives should be closed. He also places an emphasis on the importance of 'a deeper understanding of the target audience' by asserting that "a good counter- or alternative narrative will understand the specific, and often local, vulnerabilities, attitudes, behaviours, interests and media consumption habits of a particular target audience" (p.7).

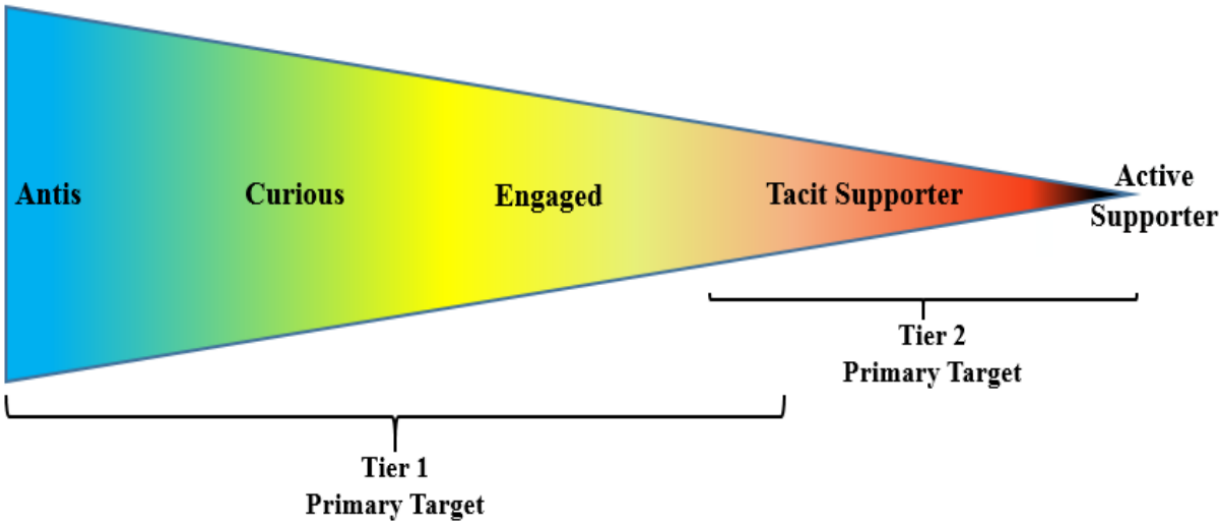


Figure 1 - Ingram's (2016) two-tiered messaging system

Schlegel (2021) approaches the issue from a different angle by arguing for the use of fictional utopian narratives and storytelling in P/CVE campaigns. Building on Ingram's 'competitive system of meaning,' which posits that "extremist narratives… may seem more appealing to some than narratives provided by mainstream society" (2017, p. 198), she asserts that narratives, with their persuasive power and capacity to influence individuals' beliefs and behaviours, can also make appealing counter- or alternative narratives for preventing or countering radicalisation processes. Furthermore, Ingram's own solution to the 'system of meaning' of extremists was to use a linkage-based approach (ibid), which entailed attacking the linkages that create a self-reinforcing cycle by using a two-tiered strategy. The first tier primarily targets those who have not yet accepted the extremist 'system of meaning' and uses a combination of negative messages that attack the linkages between extremists and their solutions, and positive messages that focus on possible solutions beyond the extremists. The second tier aims at those already within the cycle by using negative messaging as a disengagement and disruption strategy. Tailoring messages towards the correct target audience is crucial in Ingram's framework, for which he offers five categories: antis, curious, engaged, tacit supporters, and active supporters (ibid).

Other researchers place more emphasis on the role of individuals. Acknowledging the efforts of American Muslims to counter extremism, Helmus et al. (2022), for example, argues that these efforts should be expanded by amplifying the messages of credible voices such as Muslim social media influencers, artists, and playwrights. Examining a US government capacity building initiative that focuses on social media training, and Google Ideas counter-radicalization initiative which aimed to "tackle the issue of radicalization by creating a network of denounced extremists, gang-members and Neo-Nazis (referred to as 'formers')" and 'survivors' of attacks, the authors underlined the importance of being 'enablers' in such efforts instead of imposing what is right or wrong with a top-down approach (ibid, p.1,7). Similarly, arguing that "the messenger is as important as the message", Hamid (2022, p.56) asserts that P/CVE messaging should be viewed within a social ecology, for "messages from trustworthy and authoritative sources conducted person-to-person have the greatest impact on persuasion."

Braddock (2022) focuses on an inoculation-based approach, in which audiences are given a "weakened" form of an extremist narrative and told that the message seeks to manipulate them into supporting their cause. Telling the audience that they are being manipulated triggers "psychological reactance," which makes them less likely to be persuaded. In Braddock's experiment, 357 participants from the U.S. were exposed to an inoculation message (or no-inoculation control message) before reading left- or right-wing extremist propaganda. The findings show that "inoculation positively predicted reactance to extremist propaganda, which in turn, reduced intention to support the extremist group" (ibid, p.240). The findings also demonstrated that inoculation led to reduced credibility of the extremist group for the audience - which is important because this credibility is positively associated with intention to support the group. In brief, Braddock's research demonstrates good implementation of inoculation theory in the realm of countering extremism and has important ramifications for developing effective counter- or alternative narratives.

To sum up, all these approaches aim to disrupt and discredit the narrative and ideology of extremist groups, to build resilience among vulnerable populations, to foster dialogue and understanding among different groups, and to support deradicalization and rehabilitation programs.

4. Existing Guidance Documents

The increased availability and variety of global dissemination methods being leveraged by extremist groups and other malicious non-state actors has resulted in an increased effort to prevent and counter such narratives. In response to extremist and hateful messaging and propaganda, several organisations and practitioners have developed handbooks, training materials, and "how-to" guides to support civil societies, governments, practitioners, and the wider community in creating positive and alternative narratives to counteract extremist and hateful content. For this report, we selected the following as examples and will be referred to as guiding documents: 1) "The Counter Narrative Handbook"; 2) "Undermining Violent Extremist Narratives in the Middle East and North Africa: A How-to Guide"; 3) "GAMMA+"; 4) "Building Your Own Counter-Narrative Campaign on a Shoestring"; and 5) "LOUD – When local authorities and young people from nine European cities mobilise against intolerance".[9] A table summarising these reports can be found at the end of this document. The aim of this section is to present the selected guiding documents and highlight key common elements included in the development of a messaging campaign. Distinctive or contrasting aspects, as well as general challenges highlighted in these documents are also explored.

4.1. Existing Guidance Documents

The guiding principles across the documents reviewed broadly focus on the same elements: assessing the context; identifying target audience; developing clear goals and objectives; selecting a messenger; developing the message; creating a dissemination strategy and evaluating the impact of the campaign.

Context

The first identified common step is assessing the context. Generally, this step calls for the practitioners (i.e., campaign developers) to assess their context or, as mentioned in the "LOUD" handbook, carrying out a "local audit" in the selected context (2021, pp. 30-36). In this guidance document, the "audit" helps the practitioner to identify what type of harmful messaging is being spread in the context and its consumers, what the driving forces of such messages are, as well as to get a preliminary understanding of the best actors to deliver counter or alternative messages. The local audit also calls for an inventory of activities to prevent discrimination and radicalization which assesses what sort of responses are already in place, helping to enhance existing efforts and avoid duplication. According to Hedayah's "How-to Guide", assessment of the context includes understanding and examining existing "push and pull factors" or grievances that exist in the selected space at both the macro and individual levels (El Sayed, Faris and Zeiger, 2017). The authors provide a deeper understanding of such terms, identifying practical examples from the specific context, and highlighting the importance of understanding these from a local perspective. By calling on practitioners to assess or "audit" the local context, the authors highlight the importance of understanding the existing factors that should help create the foundation of any campaign. Those designing campaigns must have a deep knowledge of the issue and the appropriate methods to address it.

9 It is worth noting that these are a selection of handbooks that are familiar to the team, with no intention to evaluate or assess as the most successful. There are plenty of other existing handbooks and "how-to" guides that are not mentioned in this document.

Audience

Once the general assessment of the local context is conducted, the next common step among guiding documents is the selection of the target audience. The Radicalisation Awareness Network's (RAN) "GAMMA+" model underlines the importance of having an "in-depth understanding" of the campaign audience, as this serves as "the foundation of every intervention" (Ritzmann, Wouterse, and Verdegaal, 2019, p.6). This includes the audience's priorities, communication style and method, and how the selected audience would engage or respond to the campaign. Some of the main questions that the guiding document suggests are shown in Figure 2 below.

Audience – What do you need to know?

Can you define the key characteristics of your audience members? Can you narrow down and clearly identify their age range, gender, cultural background, 'group language' and place of residence? Do you understand why they would care about your intervention, why they would interact with you? Do you understand what could change their thinking and behaviour? Where do they get their information from? Where are their echo chambers that you need to access with your message? If your aim is to reinforce the resilience of your audience, understanding them also means understanding their vulnerabilities. What makes them vulnerable?

Figure 2: GAMMA+ model – Audience (Ritzmann et al., 2019, p.7)

An EU-funded project "One Brave Thing" similarly outlines a number of questions in the "Building Your Own Counter-Narrative Campaign on a Shoestring" (n.d, p.5) guiding document:

- Where does your audience spend their time (both online and offline)?
- What do they value, or what do they see as morally good?
- How do they communicate, or how do they use language?
- Who do they respect, or who else influences them?
- Which campaigns are they already involved in?

The authors encourage campaign developers to think beyond the typical audience (i.e., extremists), but also consider broader members of the community such as youth, community leaders, authority figures, and social media followers of certain influencers (ibid, p.4).

Goals and objectives

While it may be mentioned at different stages of the campaign development, the guiding documents call for the need to have goals and objectives that are SMART – Specific, Measurable, Acceptable, Realistic, and Temporally defined (Morales, De La Torre, and Bourgeois, 2021). Using a concrete case study, Hedayah's guiding document outlines the steps to set goals and objectives. This includes "assessing push and pull factors", "identifying target audience", "identify explicit/implicit violent extremist narrative being countered", and "set clear goals and objectives for the counter-narrative" (El Sayed et al., 2017, p.16). To a varying level of detail, the guiding documents encourage campaign developers to establish a monitoring and evaluation plan, which should help assess their campaign's success based on the agreed goals and objectives. This should collate all information gathered for the campaign development, assess available resources, and identify all activities carried out to help assess how the desired goals and objectives have contributed to the intended impact. If required, the goals and objectives may be re-evaluated based on various factors relating to the campaign project.

Message Framing

While an obvious point, the message is one of the most crucial and sensitive elements of a campaign. "The Counter Narrative Handbook" describes messages as stories "with a purpose," which "do not lecture the audience - they offer something to think about and reflect on" (Tuck and Silverman, 2016, p.9). The document goes on to give examples of what type of messages can be impactful in preventing or countering harmful narratives. It also underlines the importance of developing messages with care and awareness of possible side effects, ensuring that the message does not create an unintended impact and antagonise the audience. In the "Building Your Own Counter-Narrative Campaign on a Shoestring" project, the authors suggest deciding "what you want the reaction of your audience to be" (n.d, p.6), indicating a need to know the intended output of your message/campaign/story, which relates to the overall campaign goals and objectives. Whether explicitly mentioned or implicitly indicated, the guiding documents demonstrate how all the elements are interrelated. With the selection of a certain audience and messenger, campaign designers are already creating a foundation for what the message should be. If the message and the messenger do not resonate with the selected audience in the specific context, then the positive impact is likely to be minimal.

Messenger

Another critical element mentioned in the guiding documents is the campaign messenger – the individual(s) who will deliver the message to the identified audience, seeking to change their attitudes and/or behaviours. As stated in the Institute for Strategic Dialogue's (ISD) guiding document, "[t]he lyrics might be powerful but the singer also needs a good voice" (Tuck and Silverman, 2016, p.11). This underlines the importance of carefully selecting messengers who are considered legitimate and credible actors in the eyes of the selected audience. Understanding the intended audience is helpful for identifying the most effective messenger to deliver the message with the greatest possibility of a successful impact. ISD's "The Counter Narrative Handbook" provides various studies as examples of audience-relevant messengers, as well as highlighting the credibility messengers may leverage in delivering certain messages based on their background. Hedayah's "How-to Guide" (El Sayed et al., 2017) provides several questions that may help an individual identify their messenger (see Figure 3 below). Furthermore, the document also provides an example of common campaign messengers, highlighting benefits of each (ibid). The authors of the "GAMMA+" document similarly reiterate the need to understand your audience before selecting a messenger (Ritzmann et al., 2019).

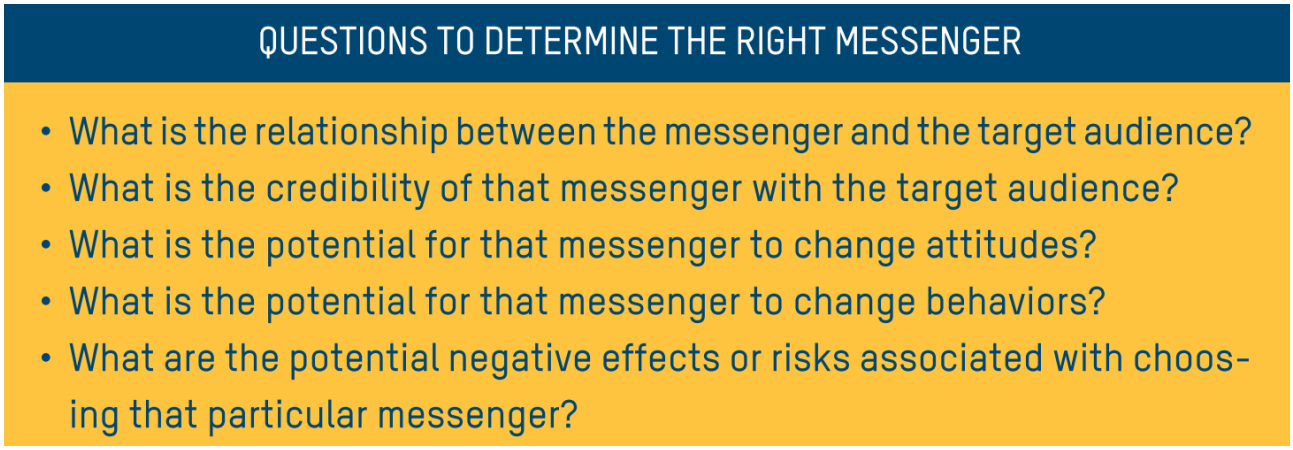


Figure 3: Hedayah's list of question to determine the right messenger (El Sayed et al., 2017, p.17)

Strategy and evaluation

Final common elements shared by successful messaging campaigns relate to strategy and evaluation. The first component – strategy – looks at the best approach to disseminate the message most efficiently. Hedayah's guiding document outlines several recommendations for the "strategic dissemination of messages" shown in Figure 4 below (El Sayed et al., 2017, p.28). In the "LOUD" document, the authors call for campaign developers to first "meet your audience where they are", leveraging "appropriate channels" and adapt the message to the selected online or offline medium (videos, posters, audios, articles, etc.) (Morales et al., 2021, p.43).

Figure 4: Hedayah's recommendation on message dissemination strategy (El Sayed et al., 2017, p.28)

The selected dissemination strategy will also impact the final evaluation of the campaign. Though briefly touched upon in relation to goals and objectives, evaluation relates to every element of the process. All the guiding documents provide some level of description of the evaluation process, encouraging campaign developers to develop messages in such a manner that will allow them to later assess its impact through development of measurable indicators linked to goals and objectives. The process of campaign monitoring and evaluation will be discussed in detail in Section 6 (see below).

4.2. Distinctive elements

Analysis of the guiding documents identifies several distinctions in the advice provided for implementing alternative messaging campaigns. The first of these distinctions can be found in the "LOUD" guide which recommends the explicit application of a substantial and formal pre-campaign 'audit' (Morales et al., 2021, pp. 30-36). Whilst the other guiding documents also recognise the importance of properly understanding the relevant context, the "LOUD" authors go one step further to suggest the implementation of an elaborate audit of the target audience and locality through the extensive pre-campaign collection of both qualitative and quantitative data regarding existing hate and counter-hate conditions within the focus locality (ibid, p.24). Building on the 'context assessments' suggested in other guides, this seeks to ensure the greatest message and messenger resonance through the delivery of hyper-local alternative narratives. Additionally, the "LOUD" guide is distinct in stressing the need to train local alternative narrative messengers (ibid, pp.36-38). While some guides propose digital literacy training for messengers, the "LOUD" document suggests that alternative narrative messengers be trained extensively in the dynamics of hateful content dissemination, the 'real world situation' of the focus locality (using the expansive data provided by the audit), and the sustainable delivery of interactive messaging campaigns (ibid, p.37).

The prescription of a local focus is shared by Hedayah's how-to guide. While the "LOUD" guide develops a proposal for individual city-based alternative narrative campaigns, the Hedayah guide builds upon several other regional handbooks that are referenced in the introduction (El Sayed et al., 2017, pp.1-2). These guides seek to develop more regionally applicable frameworks that take into account both the dominant hateful narratives of each region, and the relevant actions that can be taken to deliver a more appropriate alternative-narrative campaign.

A further distinction of note can be found in the "Counter-Narrative Handbook" (Tuck and Silverman, 2016). The authors make explicit reference to employing the use of targeted social media advertising for counter-narrative campaigns (ibid, pp.30-35). Such advice is predicated on a belief – shared by the other guides – that campaigns can only be effective if they successfully reach their target audience in the places where this audience congregates. Given the ease, efficiency, and customizability of targeted social media advertising, it is a valuable recommendation that could aid in the effective delivery of alternative messages to the most appropriate individuals, should the campaign's budget allow.

Despite the proliferation of online hate communities – and the growing need to counter them - a final distinction of note acknowledges the continued importance of offline action to complement online campaigns. The "GAMMA+" document recognises the need for counter-narrative campaigns to be repeatedly linked to offline behaviour (Ritzmann et al., 2019, pp.10-11) and gives extensive consideration of the role played by parallel offline functionality. While the au-

thors acknowledge this goal is not necessarily representative of all alternative messaging campaigns, they simultaneously concede that the most successful online campaigns will be those predicated on instigating a co-ordinated offline behavioural change in the target audience. They accept such desired offline behaviour to be unique to each campaign but highlight that more effective counter-narrative campaigns will be those that do not simply encourage the target audience to question hateful narratives, but instead demonstrate real-world behavioural change. It is this extensive reference to supplementary offline engagement that emphasises the dialogic and reciprocal nature of counter-narrative campaigns whereby audience and practitioner actively engage one-another (Ritzmann et al., 2019, p.11).

4.3. Challenges and common pitfalls

The guiding documents also outline some common challenges and pitfalls faced in alternative narrative campaign development. One of the challenges mentioned is the potential harms associated with mishandling certain elements of the process. "GAMMA+" underlines the dangers of lacking sufficient information and understanding of the selected audience. Underestimating the "complexity of an audience" may lead to various issues, such as exposing an indifferent audience to extremist propaganda (ibid, p.7). Poor selection of the campaign message or messenger may also have an opposite effect to that intended, antagonising the audience in the process (Tuck and Silverman, 2016). The RAN document provides a "How Can I Stay Safe Online" toolbox to help provide security advice throughout the process of delivering interventions. Addressing evaluation, a common pitfall is withholding monitoring and evaluation processes until the end, rather than including them from the inception phase of the campaign development (Ritzmann et al., 2019). In "GAMMA+", the authors highlight the consistent need to have more campaign testing to support the learning process and improve the successful impact of such interventions (ibid).

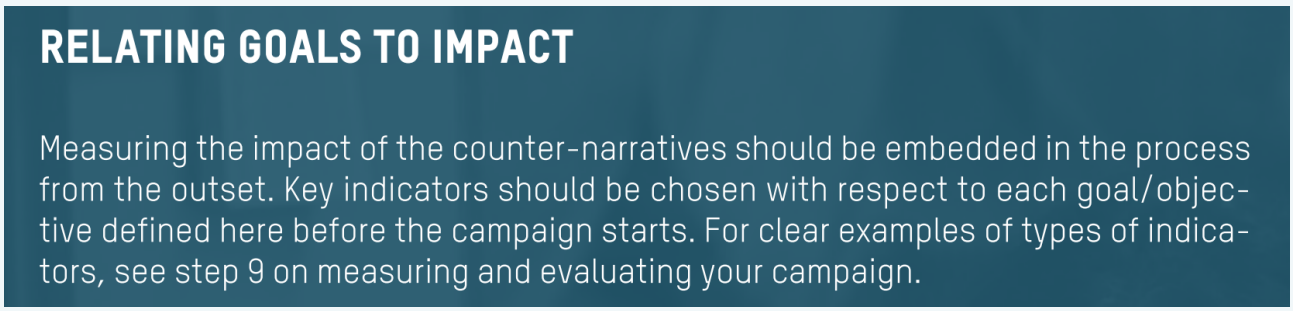


Figure 5: Hedayah's how-to guide "Undermining Violent Extremism Narratives in The Middle East and North Africa" (El Sayed et al., 2017, p.18)

As shown above, most of the documents share various common elements and identify common challenges, but also demonstrate some substantive distinctions. There is a need to ensure that such guidance materials are updated and adapted to the evolving nature of communication mediums and techniques. In the next section we will explore audio-visual considerations of such counter-narratives.

5. Considerations for P/CVE Short-Form Video Content

So far, this report has focused on the key lessons that have been learned in developing effective strategic communications for P/CVE. However, short-form videos have specific considerations that may set the medium apart from other forms for delivering prosocial messaging. As such, this section outlines five factors worthy of consideration for content creators: format; production value; humour and tone; sound; and maintaining safety.

5.1. Format

An initial consideration when devising P/CVE content is to decide upon a format that will resonate with the target audience. Narrative-focused content that draws upon the first-hand accounts of individuals impacted by a particular issue provides an emotionally compelling, relatable, and ultimately persuasive means for delivering pro-social messaging (Romney and Johnson, 2020, p.106). The life-story format, whereby former members (those who have previously joined violent groups) or individuals who have been directly impacted by extremism describe their experiences, has become a staple method of countering extremism - the assumption being that those who have 'been there and seen it' can most effectively undercut extremist narratives (Rrustemi, 2020, p.2; Braddock and Morrison, 2020, p.483).

Content featuring storytellers with similar backgrounds or characteristics to the audience has been shown to be particularly engrossing, including where this content is aimed at countering extremist narratives (Green, 2004; Copeland, 2019; Monaci, 2020). The narrator and central character of the "Abdullah-X" series of counter-narrative animations, for example, appears to have been designed to resonate with a young audience, with chunky headphones around his neck, graffiti on the backgrounds behind him, and a thick London accent[10]. Other campaigns

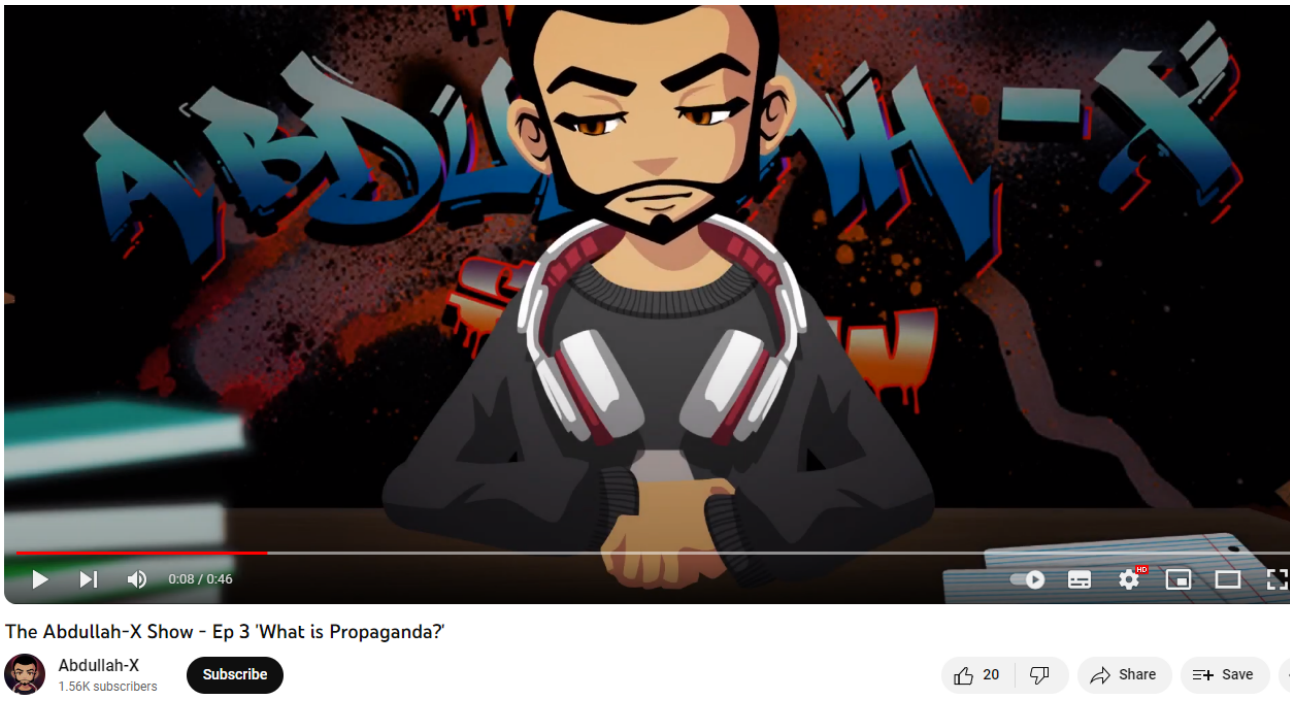


Figure 6: A screenshot showing the main character in the Abdullah-X YouTube video series.

10 See https://www.youtube.com/user/abdullahx.

deliberately keep the appearance of their central characters vague, in an effort to encourage viewers to directly project themselves into these roles (Ali et al., 2020).

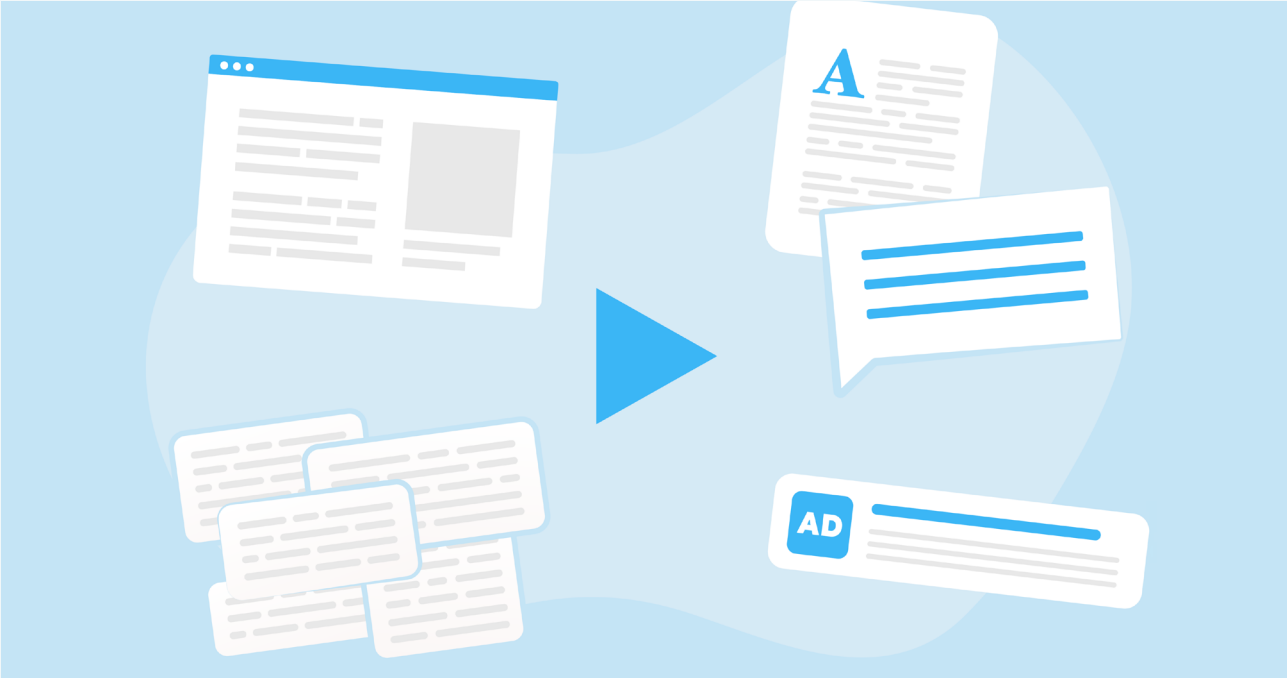
Significantly, research finds that the inclusion of first-hand accounts remains an effective approach for delivering prosocial messaging through short-form video content (Shriver-Rice et al., 2022, p.71). However, the short duration of these videos means such content must be carefully tailored. For example, one video series attempted to use footage of individuals discussing conspiracy theories to counter the narratives that underpin them. However, an evaluation found that most viewers did not watch the videos until the end and skipped past the twist towards critical thinking. Those inclined to conspiratorial thinking therefore ended up only consuming material that strengthened their existing views (Ritzmann et al., 2019). Successful short-form content creators also understand the importance of an impactful opening 'hook' to beat the skip button (Abboud, 2021, p.629). Viewers commonly refuse to watch more than five seconds of a short-form video that did not capture their attention by being too slow, boring, and repetitive, or – in the case of pro-social messaging – too emotional (Shriver-Rice et al., 2022).

Successful pro-social, short-form videos generally appear to quickly convey clear and straightforward messages (Rosato, 2021, p.24; Rrustemi, 2020, p.7). The key point should be presented immediately in an identifiable and engaging way (for example, using a particularly salient quote from an individual's story) which can then be followed with a more detailed explanation once the users' attention has been captured (Shriver-Rice et al., 2022). Whilst short-form video may not immediately appear to present opportunities to convey detailed information, research highlights that comprehensive captions provide effective opportunities for engagement. Significantly, the length of caption text positively correlates with engagement, suggesting that once a viewer's interest has been captured by a video they will proactively undertake significant reading around it (Cuevas-Molano, Sánchez-Cid, and Gordo-Molina, 2022; Kuncoro, 2022, p.49). Captions also present the opportunity to reinforce other aspects of counter-narrative delivery – for example, where actors are used to portray real-life stories, a disclaimer in the video's description can highlight that the content is based on a true story (Rrustemi, 2020, p.8).

5.2. Production Value

Research highlights the increasing audio-visual sophistication of propaganda produced by contemporary extremist groups. Included are 'high production' techniques such as the use of graphics, special effects, colours, suspenseful or intense music, multiple cuts and edits, unusual lighting and camera angles, zooms, and close-ups. To effectively counter terrorist groups' engrossing propaganda, it is necessary to provide similarly immersive content (Braddock and Horgan, 2016, p.481). Researchers have argued, for example, that by taking advantage of TikTok's affordances, (e.g., adding filters, effects or stickers or interacting with other videos through options such as "green screen," "paste from," or video responses to comments from other accounts), users can effectively participate in social and political debates (Civila, del-Rio and Aguaded, 2023, p.4). However, it also appears that engagement can be generated without the use of sophisticated audio-visual techniques. Whilst analysis of anti-smoking campaigns on YouTube has found that viewers interacted more with videos with higher production values (Paek, Kim, and Hove, 2010), an evaluation of existing counternarrative videos finds that audio-visual complexity of these campaigns varies considerably. There is no evidence that videos at the higher end of the production value spectrum are more effective (Monaci, 2020).

Short-form video content is also created largely by amateur performers in a less-polished manner than other social media platforms. This imperfection often closes the distance between the producers and consumers of content elevating engagement (Su, Baker, Doyle, and Yan, 2020, pp. 441-442). Some studies also suggest that the presence of too many audio-visual elements and techniques decreases engagement as viewers struggle to make sense of overly complex videos in such a short duration (Kuncoro, 2022, p.48). A less-cultivated and spontaneous appearance may therefore be beneficial for short-form counter-narrative content, helping videos to appear authentic and relatable whilst remaining easily consumable. Content should, however, be optimised for viewing on mobile devices - research finds that short-form videos displayed in vertical screen mode received more engagement (likes, comments, and shares) than those in horizontal mode (Zhuang, 2022, p.33).



5.3. Humour and Tone

The tone of counter-narrative content has often been identified as a difficult balance to effectively strike. Many campaigns have been criticised for having a 'preachy' or condescending tone (Tierney, 2017). Irony, satire, and parody are central features of some counter-narrative content, with ridicule thought to be a potentially effective way to discredit extremist narratives without appearing patronising (Lee, 2019). Spoofs have proven an engaging and enjoyable way for social media creators to call attention to, and dismantle, a variety of harmful misrepresentations and invite discussion and reflection on important issues (Maagaard and Lundhold, 2018, p.134). However, successfully using humour with such a sensitive topic as extremism can be extremely difficult and at times has attracted controversy and criticism (Rrustemi, 2020, p.3). Focus group respondents evaluating one counter-narrative campaign universally agreed that humour was ineffective and even counterproductive (Mazza, Monaci, and Taddeo, 2017, p.16). More broadly, the use of humour in short-form videos promoting social and political issues has been found to be polarising, provoking either a strongly positive or negative response from viewers. Creators are often labelled as trying too hard or making light of serious issues where jokes do not immediately land.

Humour and satire should only be used where the humour of the target audience is well known (Shriver-Rice et al., 2022, p.75). It is also crucial to consider the target of ridicule. For example, attempting to make a particular terrorist leader look foolish might successfully undercut their appeal and has a relatively limited risk of producing a counterproductive effect. However, psycho-social research suggests that humour is likely to elicit a defiant response if the audience associates with the butt of the joke, even in a limited sense. Ridiculing would-be recruits or those who display some interest or sympathy toward extremists' narratives is therefore likely to be counterproductive if the target audience includes such at-risk individuals (Beutel et al., 2016).

5.4. Sound

The use of sound is thought to be particularly important for the success of short-form video content. Videos with sound generate significantly higher engagement in terms of likes and shares (Cuevas-Molano et al., 2021). Whilst subtitles are also generally thought to increase engagement, focus groups responded negatively to videos that relied solely on captions to explain their message, respondents stressing that they want to be told what is happening rather than reading it on screen. By contrast videos that included voice overs were viewed as more authentic, personal, and persuasive (Shriver-Rice et al., 2022, p.72).

The use of popular or 'trending' music or audio is thought to increase the promotion of content on short-form video platforms and can make content familiar to viewers – something that extremists are known to exploit (Weimann and Masri, 2020, p.6). Studies of short-form video find a direct relation between soundtrack and viewer engagement. Research with focus groups shown content promoting social causes found that poor musical choices were a key feature of uncompelling or uninteresting videos with participants more likely to recall music they did not like, rather than that which they thought was particularly good. All respondents agreed that annoying, overly dramatic, or overly happy and/or upbeat music was off-putting in videos focusing on serious issues (Shriver-Rice et al, 2022, p.70). On some social media platforms, such as TikTok, it is also possible to attach trending audio to a video but to turn the sound down, content therefore benefitting from the algorithmic boost of using a trending audio without interfering with the video's core message.

5.5. Safety Considerations

As noted, life story formats have become a prominent feature of counter-narrative content. Where those telling their stories have been involved or are at risk of being drawn into extremism maintaining anonymity is crucial – something important to consider when filming. For example, creators might consider if the location or background stand out? Does it include landmarks or other features that could be recognised? Even seemingly benign features such road markings

or prominent background sounds (such as heavy air traffic) might reveal more than intended. The need to maintain anonymity, however, does not mean interviews cannot be captured in interesting or engaging ways. Close-up shots of the subject's hands or lower face or those taken from behind or in silhouette all provide visually engaging ways to draw attention to their words whilst keeping their identity hidden. Rather than using potentially identifiable urban locations, natural settings such as parks, forests, or open outdoor spaces offer interesting but generic backdrops (Rrustemi, 2020, p.9).

6. Monitoring, Measurement, and Evaluation in CVE Messaging

Despite the widespread use of strategic communications in P/CVE, there is a relatively low evidence base to understand their efficacy. Ferguson argues that many campaigns are not conducting evidence-based interventions or evaluations: "Today there are numerous NGOs pursuing [P/]CVE counter-narrative projects, and many are doing so without research-driven position papers, an evidence base, or even a theory of change that sets out measurable objectives" (Ferguson, 2016, p. 11). Mattei and Zeiger (2018) highlight several reasons as to why it is so difficult to evaluate these types of programs, including the fact that interventions may take a long time to work; there can be difficulty demonstrating causality; the persistence of practical challenges to data collection; and a systemic lack of resources.

Systematic reviews of P/CVE campaigns demonstrate this lack of robust evaluation. In a meta-review of 139 communications campaigns, Jones (2020) found that none of the campaigns demonstrated effective interventions, with only 22 being "potentially effective" and the rest providing mixed, ineffective, or inconclusive results. Only 41 of these campaigns had what was deemed to be high-quality evidence with which to make such a judgement. In total, only 9 of the "potentially effective" campaigns utilised high-quality evidence. Similarly, in a systematic review of all P/CVE interventions conducted by Hassan et al. (2021), it was found that of 11,836 studies identified, only 31 met the appropriate criteria and methodological quality to be included. Of these 31 studies, only three messaging campaigns were identified. They note that these three do show mostly positive results, though none actually measured the effect of viewing a campaign message upon extreme attitudes or behaviours, which the authors believe limits the positive conclusions that can be drawn.

There are some evaluation techniques that should be considered as part of any P/CVE messaging intervention. As noted above by Mattei and Zeiger (2018), it is fundamentally difficult to isolate the effects of the message from other factors. The implementation of two social science methods – base rates/baselines and control groups – can help to do this, although it should be noted that both methods are imperfect. Using base rates requires data to be collected prior to any intervention implementation, so that this data can be compared against the post-intervention results. For example, if a message is designed to reduce hostility toward a marginalised community, one might conduct a survey prior to the intervention to gauge their feelings, either within the general population (base rate) or the participants of the study (baseline), then again afterwards to assess whether the intervention has worked. This is vital in overcoming the "base rate fallacy" whereby people ascribe greater weight to information specific to the event, whilst failing to pay attention to how prevalent this sentiment is within the general population (Bar-Hillel, 1980). In our example, if one were to only test hostility to the community after the

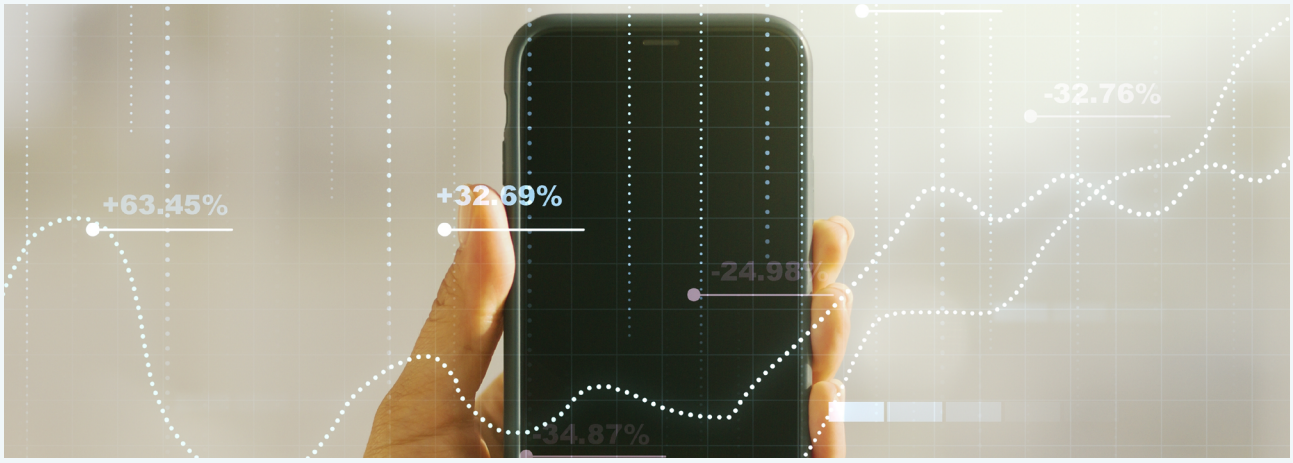
intervention and found that most people held "somewhat negative" sentiments, this information is redundant without comparing it against, either a base rate of the general population, or a baseline of the participants. It could be that the intervention has, in fact, caused them to move from "very negative" sentiment, which would be a success, or "positive" sentiment, in which case the intervention would have backfired.

Another technique is to employ a control group, who do not receive the intervention condition, to facilitate comparison against the individuals that do receive one. In principle, this should demonstrate whether the intervention has affected the dependent variable: in this case, using the above example – sentiment towards a marginalised community (Bryman, 2015). This has been employed in several ways in P/CVE messaging. In their laboratory-style experiments on counter-narratives, both Carthy and Sarma (2021) and Braddock (2022) leave a group of participants with no (or a neutral) narrative for comparison against those that receive an intervention. In their study of counterspeech on Facebook, Saltman, Kooti, and Vockery (2021) wanted to assess whether an intervention affected the audiences' continued engagement with, and consumption of, extremist content online. Therefore, they randomly divided their participants into two groups, before showing one group the counter-messages, and the other nothing at all. They then observed the average rate of extremist content consumption in both groups and compared them.

Another method of evaluation that could involve either baselines and control groups (or both) is conducting a linguistic analysis. These approaches have been used to analyse extremist behaviours online but have seemingly not been deployed to analyse the efficacy of interventions. Sentiment analysis is a popular method which quantifies words into positive and negative scores to assess or detect radical sentiment (Bermingham et al., 2009; Scrivens, Davies, and Frank, 2017; Scrivens, 2020). Research has also utilised a Corpus-Assisted Discourse Studies (CADS) approach which combines both qualitative and quantitative textual analysis (Lorenzo-Dus and Macdonald, 2018; Nouri and Lorenzo-Dus, 2019). Reynolds and Tuck (2016) propose linguistic methods as a potential method for evaluating P/CVE strategic communications, noting that it can categorise large scale online content. Theoretically, prior to conducting an intervention, one could analyse the existing language on an online platform, conduct the intervention, then conduct another round of analysis, testing the changes – i.e., comparing to the baseline. Alternatively, one could split participants into two groups, and only provide one condition with the intervention material and then compare the language expressed with those of the condition who did not receive intervention material – i.e., a control group.

Another avenue for evaluation is to categorise messages (or wider campaigns) and compare them against each other. Utilising Ingram's linkage-based framework and 'system of meaning' concept and analysing 10 social media multi-message campaigns, Whittaker and Elsayed (2019) offered five different 'ideal types,' which entail: i) aggressive type that highlights an opponent's actions (i.e., atrocities); ii) argumentative type, a defensive strategy that highlights why an extremist message is wrong; iii) assertive type, which combines positivity and negativity, while prioritizing the former; iv) know it all type, which relies on a cost-benefit analysis by highlighting the problem and solutions; and v) identity-builder type, which resembles the know-it all type but differs from it given that the former focuses on cost-benefit analysis, while the latter is focused "on instilling a sense of identity in its audience" (2019, pp. 27-18). They argue that such a categorisation can aid evaluation by assessing which ideal types resonate with specific audiences.

Caution should be exercised in interpreting social media data when evaluating P/CVE communications. There may be a temptation to use metrics such as impressions, reach, and views to assess whether a message or campaign is successful. However, these are typically known as "vanity metrics" because though they may sound impressive, they do little to explain whether the message has been successful in creating tangible changes to behaviour or attitude (Reynolds and Tuck, 2016). This has been an enduring problem in the online P/CVE space, which has, in part, led to the insufficient evidence base laid out above (Zamir, 2022). There are a range of metrics which offer a greater insight into whether audience behaviours have changed, such as Click Through Rate[11], which has been utilised in previous messaging interventions (Moonshot, 2016; Saltman, Kooti, and Vockery, 2021), but these should also be treated with caution as they do not capture whether an individual has altered their beliefs or is, as a result, less likely to engage in violence.



7. Conclusion

The rise of short form videos seems to show no signs of abating. This report, which accompanies a non-technical guidance document for content creators, has offered a background to those that seek to leverage this medium to counter extremism. It is by no means exhaustive but represents many of the important lessons that have been learned in the field of P/CVE over the past two decades. This includes knowledge on the broad topic of persuasion; highlighting the difficulty in changing individuals' behaviors and attitudes, particularly when it comes to their "sacred values." This is followed by discussions of strategic communications designed to dissuade people from extremism, which includes existing guidance documents created by practitioners. This section shows that there are a range of different approaches by practitioners, although there are several important guiding principles that should be adhered to such as the context; audience; goals and objectives; message framing; messenger; and strategy and evaluation. The report moves on to considerations that are specific to succeeding at the short-form audio/visual medium, including the format; production value; humor and tone; sound, and the importance of maintaining anonymity when working in this field. The report concludes with ways in which such messages can be evaluated for efficacy in future.

11 This a ratio showing how often people click on an online advertisement or intervention compared to how many have seen it. It is calculated by first measuring how many people the intervention reached ("impressions") and then dividing it by the number of times people clicked on it. For example, if an intervention has 100 impressions and is clicked on 5 times, it has a click through rate of 5%.

References

Abboud, D. (2021). Factors driving higher engagement rates of digital video advertising. Journal of Architecture, Arts and Humanistic Sciences, 6(28), pp. 620–631. DOI: 10.21608/mjaf.2020.20681.1414.

Adena, M., Enikolopov, R., Petrova, M., Santarosa, V., and Zhuravskaya E. (2015). Radio and the Rise of the Nazis in Prewar Germany. The Quarterly Journal of Economics, 130(4), pp. 1885–1939. DOI: 10.1093/qje/qjv030.

Ali, R., Šibljaković, D., Lippe, F., Neuberg, U., and Neuburg, F. (2020) "You're against Dawla, but you're listening to their Nasheeds?' Appropriating jihadi audiovisualities in the online streetwork project Jamal-al-Khatib — My Path!" In: C. Günther and S. Pfeifer (editors). Jihadi audiovisuality and its entanglements: Meanings, aesthetics, appropriations. Edinburgh: Edinburgh University Press.

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. Acta Psychologica, 44(3), pp. 211–233. DOI: 10.1016/0001-6918(80)90046-3.

Bermingham, A., Conway, M., McInerney, L., O'Hare, N., and Smeaton, A. F. (2009). Combining social network analysis and sentiment analysis to explore the potential for online radicalization. 2009 International Conference on Advances in Social Network Analysis and Mining, Athens, Greece, pp. 231–236. DOI: 10.1109/ASONAM.2009.31.

Beutel, A., Weine, S., Saeed, A., Spahic Mihajlovic, A., Stone, A., Oakley Beahrs, J., and Shanfield, B. (2016). Field Principles for Countering and Displacing Extremist Narratives. Journal of Terrorism Research, 7(3), pp. 35–49. Retrieved from: https://cvir.st-andrews.ac.uk/article/10.15664/jtr.1220/.

Braddock, K. (2022). Vaccinating Against Hate: Using Attitudinal Inoculation to Confer Resistance to Persuasion by Extremist Propaganda. Terrorism and Political Violence, 34(2), pp. 240–262. DOI: 10.1080/09546553.2019.1693370.

Braddock, K, and Horgan, J. (2016). Towards a Guide for Constructing and Disseminating Counternarratives to Reduce Support for Terrorism, Studies in Conflict & Terrorism, Vol. 39(5); pp 381-404.

Braddock, K. and Morrison, J. (2020). Cultivating Trust and Perceptions of Source Credibility in Online Counternarratives Intended to Reduce Support for Terrorism. Studies in Conflict & Terrorism, 43(6), pp. 468–492. DOI: 10.1080/1057610X.2018.1452728

Bryman, A. (2015). Social Research Methods. Oxford: Oxford University Press.

Carmen, A. (2020). TikTok reaches 2 billion downloads, The Verge, April 29. https://www.theverge.com/2020/4/29/21241788/tiktok-app-download-numbers-update-2-billion-users.

Carthy, S. L. and Sarma, K. M. (2021). Countering Terrorist Narratives: Assessing the Efficacy and Mechanisms of Change in Counter-narrative Strategies. Terrorism and Political Violence, 35(3), pp. 569–593. DOI: 10.1080/09546553.2021.1962308.

Civila, S., del Rio, M.B., and Aguaded, I., (2023). Social Media and Otherness: The Case of #Islamterrorism

on TikTok. Politics and Governance, 11(2), pp. 1-13. DOI: https://doi.org/10.17645/pag.v11i2.6299

Copeland, S. (2019). Telling Stories of Terrorism: A framework for applying narrative approaches to the study of militant's self-accounts. Behavioral Sciences of Terrorism and Political Aggression, 11(3), pp.232–253. DOI: 10.1080/19434472.2018.1525417.

Cuevas-Molano, E., Sánchez-Cid, M., and Gordo-Molina, V. (2022). Brand Strategy and Content Management on Instagram: Scheduling and message length as factors to improve engagement. Comunicacion y Sociedad, 35(2), pp.71–87. DOI: 10.15581/003.35.2.71-87.

Divon, T. & Ebbrecht-Hartmann, T., (2022). "#JewishTikTok" in Boffone, T. [Ed.] TikTok Cultures in the United States, London: Routledge, pp. 47-58.

Do One Brave Thing. (n.d.). Building Your Own Counter-Narrative Campaign on a Shoestring. Retrieved from: https://onebravething.eu/wp-content/uploads/2020/05/BRAVE_D5.4_ENG.pdf.

Downs, J. (2018). Increasing our Investment in Creators for Change, YouTube Blog, 24th January. https://blog.youtube/creator-and-artist-stories/increasing-our-investment-in-creators/.

El Sayed, L., Faris, T., and Zeiger, S. (2017). Undermining Violent Extremist Narratives in the Middle East and North Africa: A How-To Guide. Hedayah. Retrieved from: https://hedayah.com/resources/undermining-violent-extremism-narratives-in-the-middle-east-and-north-africa-a-how-to-guide/.

Ellul, J. (2006). The Characteristics of Propaganda, in: Garth S. Jowett and Victoria O'Donnell [Eds.], Readings in Propaganda and Persuasion (London: Sage Publications, 2006)

Ferguson, K. (2016). Countering Violent Extremism Through Media and Communication Strategies: A review of the evidence. Partnership for Conflict, Crime & Security Research. Retrieved from: https://gsdrc.org/document-library/countering-violent-extremism-through-media-and-communication-strategies-a-review-of-the-evidence/.

Ginges, J. et al. (2007) Sacred Bounds on Rational Resolution of Violent Political Conflict', Proceedings of the National Academy of Sciences (Vol. 104, No. 18)pp. 7357-60.

Green, M. (2004). Transportation into Narrative Worlds: The role of prior knowledge and perceived realism. Discourse Processes, 38(2), pp. 247–266. DOI: 10.1207/s15326950dp3802_5.

Hamid, N. (2020). The Ecology of Extremists' Communications: Messaging Effectiveness, Social Environments and Individual Attributes. RUSI Journal, 165(1), pp. 54–63. DOI: 10.1080/03071847.2020.1731248.

Hamid, N. (2022). Mass Media and Persuasion: Evidence-Based Lessons for Strategic Communications in CVE. Cross-Border Conflict Evidence, Policy and Trends. Briefing Note. Retrieved from: https://icsr.info/2022/07/21/mass-media-and-persuasion-evidence-based-lessons-for-strategic-communications-in-cve/.

Hassan, G., Brouillette Alarie, S., Ousman, S., Kilinc, D., Savard, É. L., Varela, W., Lavoie, L., Fetiu, A., Harris Hogan, S., Borokhovski, E., Pickup, D., Madriaza, P., Rousseau, C., Thompson, S. K., McCoy, J., Venkatesh, V., Boivin, M., Srimathi Narayana, M., Morin, D., and the CPN PREV team. Hassan, G. et al. (2021)

'A Systematic Review on the Outcomes of Primary and Secondary Prevention Programs in the Field of Violent Radicalization', Canadian Practitioners Network for the Prevention of Radicalization and Extremist Violence. Retrieved from: https://cpnprev.ca/wp-content/uploads/2021/03/CPN-PREV-2nd-Systematic-Review-2.pdf.

Helmus, T. C., York, E., and Chalk P. (2013). Promoting Online Voices for Countering Violent Extremism. RAND Corporation. Retrieved from: https://www.rand.org/pubs/research_reports/RR130.html.

Ingram, H. J. (2017). The Strategic Logic of the "Linkage-Based" Approach to Combating Militant Islamist Propaganda: Conceptual and Empirical Foundations. The International Centre for Counter-Terrorism – The Hague, 8(6). DOI: http://dx.doi.org/10.19165/2017.1.06

Instagram. (2020). Introducing Instagram Reels, About Instagram, 5th August. https://about.instagram.com/blog/announcements/introducing-instagram-reels-announcement.

Jones, M. (2020). Through the Looking Glass: Assessing the Evidence Base for P/CVE Communications. RUSI Occasional Paper, (July). DOI: 10.1093/jiplp/jpu004.

Kaplan, J. T., Gimbel, S. I., and Harris, S. (2016). Neural Correlates of Maintaining One's Political Beliefs in the Face of Counterevidence. Scientific Reports, 6(December), pp. 1–11. DOI: 10.1038/srep39589.

Kiper, J. (2015). Toward an Anthropology of War Propaganda. Political and Legal Anthropology Review, 38(1), pp. 129–146. DOI: 10.1111/plar.12090.

Kuncoro, A. (2022) How to Improve Instagram Reels Engagement? An empirical investigation of content strategies. Lund University Student Papers. Retrieved from: https://lup.lub.lu.se/student-papers/search/publication/9084630.

Lee, B. (2019). Informal Countermessaging: The Potential and Perils of Informal Online Countermessaging. Studies in Conflict & Terrorism, 42(1-2), pp.161–177. DOI: 10.1080/1057610X.2018.1513697.

Lee, J. & Lee. J. (2023). #StopAsianHate on TikTok: Asian/American Women's Space-Making for Spearheading Counter-Narratives and Forming an Ad Hoc Asian Community, Social Media & Society, Vol. 9(1), pp.1-11.

Lorenzo-Dus, N. and Macdonald, S. (2018). 'Othering the West in the Online Jihadist Propaganda Magazines Inspire and Dabiq', Journal of Language Aggression and Conflict, 6(1), pp. 79–188. DOI: 10.1080/1057610X.2018.1513697.

Maagaard, C. and Lundholt, M. (2018). Taking Spoofs Seriously: The counter-narrative potential of spoofs as critical discourse. Journal of Media and Communication Research, 34(64), pp. 116–137. DOI: 10.7146/mediekultur.v34i64.24837.

Mattei, C. and Zeiger, S. (2018). Evaluate Your CVE Results: Projecting Your Impact. Hedayah. Retrieved from: https://hedayah.com/resources/evaluate-your-cve-results-projecting-your-impact/.

Mazza C., Monaci S., and Taddeo G. (2017). Designing a Social Media Strategy against Violent Extremism Propaganda: The #heartofdarkness Campaign. In: Digitale Medien und politisch-weltanschaulicher

Extremismus im Jugendalter Erkenntnisse aus Wissenschaft und Praxis. Arbeits- und Forschungsstelle Rechtsextremismus und Radikalisierungsprävention. Retrieved from: https://www.researchgate.net/publication/324774244_Mazza_C_Monaci_S_Taddeo_G_2017_Designing_a_Social_Media_Strategy_against_Violent_Extremism_Propaganda_The_heartofdarkness_Campaign_In_Digitale_Medien_und_politisch-weltanschaulicher_Extremismus_im_Jug.

McQuail, D. (1987). Mass communication theory: An introduction (2nd ed.) London: Sage Publications, Inc.

Mercier, H. (2020). Not Born Yesterday: The Science of Who We Trust and What We Believe (Princeton, NJ: Princeton University Press)

Monaci, S. (2020). Social Media Campaigns Against Violent Extremism: A New Approach to Evaluating Video Storytelling, International Journal of Communication, 14(24), pp. 980–1003. Retrieved from https://ijoc.org/index.php/ijoc/article/view/10430.

Moonshot. (2016). The Redirect Method. Retrieved from at: https://redirectmethod.org/.
Morales, T., De La Torre, P., and Bourgeois, N. (2021). LOUD - When Local Authorities and Young People from Nine European Cities Mobilise Against Intolerance. European Forum for Urban Security. Retrieved from: https://issuu.com/efus/docs/publication_loud_en_pages_fc84a59c04a766.

Nisa, C. F., Belanger, J. J., Schumpe, B. M., and Faller, D. G. (2019). Meta-analysis of randomised controlled trials testing behavioural interventions to promote household action on climate change. Nature Communications, 10(1). DOI: 10.1038/s41467-019-12457-2.

Nisa, C. F., Sasin, E. M., Faller, D. G., Schumpe, B. M., and Belanger, J. J. (2020). Reply to: Alternative meta-analysis of behavioural interventions to promote action on climate change yields different conclusions. Nature Communications, 11(1), pp. 10–12. DOI: 10.1038/s41467-020-17614-6.

Nouri, L. and Lorenzo-Dus, N. (2019). Investigating Reclaim Australia and Britain First's Use of Social Media: Developing a New Model of Imagined Political Communities Online. Journal for Deradicalization, 18(Spring), pp. 1–37. Retrieved from: https://journals.sfu.ca/jd/index.php/jd/article/view/183.
O'Shaughnessy, N. (2004) Politics and Propaganda: Weapons of Mass Seduction (Manchester: Manchester University Press)

Paek, H., Kim, K. and Hove, T. (2010). Content Analysis of Antismoking Videos on YouTube: Message Sensation Value, Message Appeals, and Their Relationships with Viewer Responses. Health Education Research, 25(6), pp. 1085–1099. DOI: 10.1093/her/cyq063.

Reynolds, L. and Tuck, H. (2016). 'The Counter-Narrative Monitoring & Evaluation Handbook', Institute for Strategic Dialogue. Retrieved from: https://www.isdglobal.org/wp-content/uploads/2017/06/CN-Monitoring-and-Evaluation-Handbook.pdf.

Ritzmann, A., Wouterse, L., & Verdegaal, M. (2019). Effective Narratives: Updating the GAMMA+ Model. Radicalization Awareness Network (RAN). Retrieved from: https://home-affairs.ec.europa.eu/system/files/2019-12/ran_cn_academy_creating_implementing_effective_campaigns_brussels_14-15112019_en.pdf.

Romney, M. & Johnson, G. (2020). Show Me a Story: Narrative, image, and audience engagement on sports network Instagram accounts, Information, Communication & Society, 23(1), pp. 94–109. DOI: 10.1080/1369118X.2018.1486868.

Rosato, V. (2021). Evaluating the Effectiveness of Online Counter-Messaging Campaigns against Radicalisation and Violent Extremism, Rivista Trimestrale di Scienza dell'Amministrazione, 4. DOI: 10.32049/RTSA.2021.4.07.

Rrustemi, A. (2020). Technical Guidelines for Lifestory Videos: The Case on Preventing and Countering Violent Extremism and Disinformation. Hague Centre for Strategic Studies. Retrieved from: http://www.jstor.org/stable/resrep24827.

Ruddock, A. (2000). Media Effects. In Understanding Audiences: Theory and Method. London: Sage Publications, pp. 37–71. DOI: 10.4135/9780857020178.n3.

Russell, J. (2018). Developing Counter- and Alternative Narratives Together with Local Communities. Radicalization Awareness Network (RAN). Retrieved from: https://home-affairs.ec.europa.eu/system/files_en?file=2020-09/developing_counter_and_alternative_narratives_together_with_local_communities_en.pdf.

Sachdeva, S. and Medin, D. (2009). Group Identity Salience in Sacred Value Based Cultural Conflict: An Examination of the Hindu-Muslim Identities in the Kashmir and Babri Mosque Issues, Proceedings of the 31st Annual Conference of the Cognitive Science Society (Vol. 31, 2009), pp. 3111–14

Saltman, E., Kooti, F., and Vockery, K. (2021). New Models for Deploying Counterspeech: Measuring Behavioral Change and Sentiment Analysis. Studies in Conflict & Terrorism. DOI: 10.1080/1057610X.2021.1888404.

Schlegel, L. (2021). Storytelling against Extremism: How Fiction Could Increase the Persuasive Impact of Counter-and Alternative Narratives in P/CVE. Journal for Deradicalization, 27 (Summer), pp. 193–237. Retrieved from: https://journals.sfu.ca/jd/index.php/jd/article/view/467.

Scrivens, R. (2020). Exploring Radical Right-Wing Posting Behaviors Online. Deviant Behavior, 42(11), pp. 1470-1484. DOI: 10.1080/01639625.2020.1756391.

Scrivens, R., Davies, G., and Frank, R. (2017). Searching for Signs of Extremism on the Web: An introduction to Sentiment-based Identification of Radical Authors. Behavioral Sciences of Terrorism and Political Aggression, 10(1), pp. 1–21. DOI: 10.1080/19434472.2016.1276612.

Shriver-Rice, M., Fernandes, J., Johns, L., Riopelle, C., and Vaughan, H. (2022). Young Adults' Reactions and Engagement with Short-form Videos on Sea Level Rise. Environmental Communication, 16(1), pp. 63–78. DOI: 10.1080/17524032.2021.1963800.

Su, Y., Baker, B. J., Doyle, J. P., and Yan, M. (2020). Fan Engagement in 15 Seconds: Athletes' Relationship Marketing During a Pandemic via TikTok. International Journal of Sport Communication, 13(3), pp. 436–446. DOI: 10.1123/ijsc.2020-0238.

Tierney, M. (2017). Combating Homegrown Extremism: Assessing Common Critiques and New Approach-

es for CVE in North America. Journal of Policing, Intelligence and Counter Terrorism, 12(1), pp. 66–73. DOI: 10.1080/18335330.2017.1293281.

TikTok. (nd). TikTok for Good. https://www.tiktok.com/forgood?lang=en.
TikTok. (2021) Thanks a Billion!, TikTok Newsroom, 27th September. https://newsroom.tiktok.com/en-us/1-billion-people-on-tiktok

United Nations Development Programme. (2022). Young Creators across Southeast Asia partner with UNDP, the EU and TikTok to promote tolerance and diversity. 29 August. https://www.undp.org/asia-pacific/news/young-creators-across-southeast-asia-partner-undp-eu-and-tiktok-promote-tolerance-and-diversity.

United Nations Security Council. (2017). S/2017/375. https://www.un.org/securitycouncil/ctc/sites/www.un.org.securitycouncil.ctc/files/n1712081.pdf.

van der Linden, S. and Goldberg, M. H. (2020). Alternative Meta-Analysis of Behavioral Interventions to Promote Action on Climate Change Yields Different Conclusions. Nature Communications, 11(1), pp. 10–11. DOI: 10.1038/s41467-020-17613-7.

Weimann, G. and Masri, N. (2020). Research Note: Spreading Hate on TikTok. Studies in Conflict & Terrorism, DOI: 10.1080/1057610X.2020.1780027.

Whittaker, J. and Elsayed, L. (2019). Linkages as a Lens: An Exploration of Strategic Communications in P/CVE. Journal for Deradicalization, 20 (September), pp. 1–46. Retrieved from: https://journals.sfu.ca/jd/index.php/jd/article/view/243.

Wojcicki, S. (2022). Letter from Susan: Our 2022 Priorities, YouTube Blog, 25th January. https://blog.youtube/inside-youtube/letter-susan-our-2022-priorities/.

Zamir, M. (2022). Active Strategic Communications: Measuring Impact and Audience Engagement. In: Introducing 2022 GIFCT Working Group Outputs. Global Internet Forum to Counter-Terrorism. pp. 8–52. Retrieved from: https://gifct.org/wp-content/uploads/2022/07/GIFCT-22WG-PI-Impact-1.1.pdf.

Zeiger, S. (2016). Counter-Narratives For Countering Violent Extremism (CVE) In South East Asia. Hedayah. https://hedayah.com/app/uploads/2021/09/File-2792016102253.pdf.

Zhuang, Y. (2022). What Makes Brand TikTok Videos Engaging? An Elaboration Likelihood Model Perspective. [Master Thesis, Dalhousie University]. Retrieved from: https://dalspace.library.dal.ca/bitstream/handle/10222/82123/YanchenZhuang2022.pdf?sequence=1&isAllowed=y.

This report was made possible because of a grant awarded at the Terrorism and Social Media Conference Sandpit Event in 2022. The time for two researchers (Whittaker and Yilmaz) was funded by the Welsh Government European Regional Development Fund.

Appendix – P/CVE Strategic Communications Guidance Documents

Reference	Morales, T. et. al. (2021). LOUD - When local authorities and young people from nine European cities mobilise against intolerance. European Forum for Urban Security.	El Sayed, L., Faris, T. & Zeiger, S. (2017). Undermining Violent Extremist Narratives in the Middle East and North Africa: A How-To Guide. Hedayah.	Tuck, H. & Silverman, T. (2016). The Counter Narrative Handbook, The Institute for Strategic Dialogue.	Do One Brave Thing. (n.d.). Building Your Own Counter-Narrative Campaign On a Shoestring.	Ritzmann, A., Wouterse, L., & Verdegaal, M. (2019). Effective Narratives: Updating the GAMMA+ model, Radicalisation Awareness Network.
Summary of counter-narrative construction	**1. Local audit of real-world circumstances.** **2. Set goals using the SMART method:** a. Specific b. Measurable c. Acceptable d. Realistic e. Temporally defined. **3. Training young people who will conduct campaign.** **4. Plan and create campaign.** **5. Disseminate narratives** **6. Evaluate campaign.**	**1. Understanding and assessing relevant push and pull factors.** **2. Identify the target audience** **3. Identify the explicitly or implicit violent extremist narrative(s).** **4. Set clear goal(s) and objective(s) of the campaign.** **5. Determine an effective messenger** **6. Identify the dissemination medium.** **7. Develop the content and logic of the message.** **8. Develop a strategy for dissemination.** **9. Evaluate and assess the campaign impact:** **a. Awareness** **b. Engagement** **c. Impact**	**1. Planning stage:** a. Audience b. Message c. Medium d. Messenger e.Reflecting extremist actor narratives f. Setting goals and objectives g. Funding and budgeting **2. Creating content.** **3. Running the campaign.** **4. Online advertising.** **5. Engaging audience.** **6. Evaluate campaign:** a. Awareness b. Engagement c. Impact	**1. What is the narrative to be countered?** **2. Who is the audience?** **3. What is the message?** a. Argument from reason? b. Argue from authority? c. Argue from emotion? **4. Where will the campaign be run and targeted?** **5. How will the message be spread?** **6. Why should people listen to the campaign?** **7. When has the campaign been successful?**	**GAMMA+ Method:** **1. Define goal(s).:** a. Make it measurable b. Make it small/ focused c. Make it simple and concrete d. Make it time-bound **2. Identify audience.** **3. Identify message.** **4. Identify messenger.** **5. Identify or create media to be used.** **6. Call to action.** **7. Monitoring and Evaluation of the campaign at the:** a. Before stage b. Testing stage c. During stage (monitoring) d. After stage (evaluation)

Target Audience	Hyperlocal targeting of sources of extremism, with a focus on young people as both givers and receivers in such narratives.	Focused on MENA extremist actors, from online and offline sources. As part of this initiative, there were three other similar collections conducted for Southeast Asia, East Africa, and the Radical Right.	General targeting of extremist or hateful content to be countered.	Particularly focuses on targeting people vulnerable to extremism narratives, rather than extremist actors.	A generalist guide with no specified target audience, although focused predominantly on online hate actors.
Further considerations and distinctions	Makes specific inclusion of the necessity of a pre-campaign audit of the local context. As this campaign has a more 'local' focal point, such a distinction is relevant. This does not, however, diminish the potential benefits to a thorough pre-campaign audit in a wider-focused campaign.	Based on, and related to, other localised counter-narrative campaign guides that are also freely available. Additionally includes similar recommendation to establish the current real-world circumstances, albeit perhaps more generally than through a localised audit process like that in Morales et. al. (2021). Includes more specific campaign assessment guidance.	Makes explicit reference to budgeting in that it provides guidance in accordance with the budget of the campaign. This includes making mention of the possibility of utilising social media advertising for the campaign if budget allows. Additionally, it provides guidance for engaging audiences, even if budget does not allow use of targeted advertising. Includes more specific campaign assessment guidance.	Presents a more generalist, smaller-scale guide for counter-narrative delivery. Is somewhat more focused on giving individuals the tools necessary to provide those who may be vulnerable to extremism with some initial counter-narratives. This includes acknowledging the mixing of offline and online engagement with extremist narratives.	Recognises the nexus between offline and online hate actors' actions and so, while focused on delivery of online counter-narratives, this guide specifically highlights the need for such a wider campaign to come with a clear offline 'call to action'. Additionally, this guide makes clear inclusion of visual media recommendations rather than recommending textual-heavy messages.