

The Hague, 08/03/2019

A practical guide to the first rule of CT-CVE messaging

Do violent extremists no favours

This paper was presented at the 2nd European Counter-Terrorism Centre (ECTC) Advisory Group conference, 17-18 April 2018, at Europol Headquarters, The Hague.

The views expressed are the authors' own and do not necessarily represent those of Europol

Authors: Alastair G. Reed & Haroro J. Ingram

1 Abstract

The research and strategic-policy fields have been flooded with initiatives to confront the propaganda efforts of violent extremists. This paper briefly explores three trends in the field of practice – disruption, re-direction and counter-narratives – before outlining the fundamentals of an approach to counter-terrorism or counter-violent extremism (CT-CVE) messaging that is based on undermining the strengths and exploiting the weaknesses of violent extremist propaganda strategies. Focusing particularly on the challenges facing western governments confronting militant Islamist propaganda, this paper argues that the first rule of effective CT-CVE messaging must be to do violent extremists no favors. Two key lessons should be central to this effort. First, strategic communication campaign planning can be enhanced by applying the KISMI roll-out principle: keep it simple, maximize impact. Second, messaging should focus on leveraging violent extremist say-do gaps in its various forms. This paper concludes by identifying how these principles can be applied for both strategic CVE and more operational post-incident purposes.

2 Introduction

Spurred by the impact of the so-called Islamic State’s (IS) propaganda campaign online, western governments have increasingly focused their CT-CVE policies on combating violent extremist propaganda. These policies have tended to coalesce around three key approaches: disruption, redirection and counter-messaging. These largely defensive measures are reflective of (as well as contributed to) a culture of “do no harm”¹ in the departments responsible for these efforts. While a philosophy of “do no harm” is understandable, we argue that the golden rule of CT-CVE messaging should be to *do violent extremist propagandists no favors* because it contributes

¹ This refers to the notion that the first responsibility of practitioners is to not make the problem being addressed worse. It is a ‘rule’ often used to prevent a culture and mentality of action being taken for the sake of engaging in action. It is often incorrectly attributed to the Hippocratic oath, “*primum non nocere*” (for more see R.Shmerling, ‘First, do no harm’, *Harvard Health Publishing*, <https://www.health.harvard.edu/blog/first-do-no-harm-201510138421>). It is a concept captured in President Obama’s apparent statement of “Don’t do stupid shit” (see D. Rothkopf, ‘Obama’s “Don’t do stupid shit” Foreign Policy’, *Foreign Policy* <https://foreignpolicy.com/2014/06/04/obamas-dont-do-stupid-shit-foreign-policy/>).

to a more assertive, less defensive, and competitively-oriented posture. Of course, such a posture needs to be driven by appropriate strategic and operational actions.

This paper begins by examining trends in disruption, redirection and counter-narrative approaches to confronting violent extremist propaganda. It is within this context that we argue that these efforts can be further enhanced by deploying two approaches that support the “do violent extremist propagandists no favors” principle. The first refers to the synchronization of short and long form messages to help maximize the reach and impact of a strategic communications campaign (“Keep It Simple, Maximise Impact”, KISMI). The second calls for CT-CVE strategic communications to prioritize messages that focus on the disparities between what violent extremists say in their messaging and what they do. By exploiting this say-do gap, CT-CVE strategic communications undermine the credibility of violent extremists and increase dissonance with their sympathizer and supporter communities.

3 Disruption: Supply-End Targeting

The broad purpose of disruption is to eliminate the problem at source by reducing the supply of, and hence access to, extremist propaganda. Europol’s EU Internet Referral Unit (EU IRU), has led the way in this approach.² The EU IRU works by monitoring and then flagging extremist content online, and sharing this information with partners, such as internet providers and social media companies, so they can remove the offending content. Rather than seeking the removal of content by legal means - which can be problematic due to jurisdictions and definitions, not to mention time consuming - the EU IRU flags content to service providers which breaches their own terms and conditions.

Disruption has an important function by targeting the supply-end of extremist propaganda but, naturally, such an approach is limited. Critics have argued that disruption engages in an impossible-to-win game of ‘whack-a-mole’, as deleted social media accounts or extremist content simply re-appear later or elsewhere online. However, such criticisms are unfair as several studies have demonstrated that suspension and

² The EU IRU which was based on an initiative originally set up by the UK government which created the Counterterrorism Internet Referral Unit (CTIRU) in 2010.

suppression of, e.g., suspected Twitter accounts led to reductions in the activity and reach of violent extremists online.³

Whilst disruption has been useful at reducing violent extremist activities on major platforms like Facebook and Twitter, this has led to displacement to smaller platforms. On the one hand, displacement to closed platforms like Telegram allows extremists to still communicate with each other, but makes it much harder to broadcast material to a wider audience.⁴ On the other hand, smaller/newer platforms may not have the resources of their bigger cousins to confront violent extremist activity. As one expert summed up, “these propaganda machines are far too massive, sophisticated, and adaptive for isolated approaches across the tech industry—regardless of how effective one company’s efforts at content removal has been.”⁵

The first problem that a takedown strategy needs to confront – i.e. who or what to takedown – is an ongoing and constantly evolving concern. When targeting focused largely on IS or al-Qa’ida, it was a relatively uncontroversial approach. However, as targeting moved beyond IS, consensus frayed on who else (e.g. extreme right-wing) and what else (e.g. news reporting or anti-extremist messaging) should be targeted. As Google points out, judging what material is “extremist” is complex: “This can be challenging: a video of a terrorist attack may be informative news reporting if broadcast by the BBC, or glorification of violence if uploaded in a different context by a different user.”⁶ This lack of clarity can result in both missing extremist content, but also the overzealous takedown of content that crosses the line into inference in free speech. For example, Facebook have been criticized for being too slow to take down

³ J.M. Berger, “Making CVE Work: A Focused Approach Based on Process Disruption,” The International Centre for Counter-Terrorism – The Hague 7, no. 5 (2016), <https://www.icct.nl/wp-content/uploads/2016/05/J.-M.-Berger-Making-CVE-Work-A-Focused-Approach-Based-on-Process-Disruption.pdf>; J.M. Berger and Heather Perez, “The Islamic State’s Diminishing Returns on Twitter: How suspensions are limiting the social networks of English-speaking ISIS supporters,” GW Program on Extremism Occasional Paper, February, 2016, <https://extremism.gwu.edu/sites/extremism.gwu.edu/files/downloads/JMB%20Diminishing%20Returns.pdf>; Daniel Grinnell, Stuart Macdonald and David Mair, “The response of, and on, Twitter to the release of Dabiq Issue 15,” Europol, 1 May 2017, <https://www.europol.europa.eu/publications-documents/response-of-and-twitter-to-release-of-dabiq-issue-15>.

⁴ Watkin, Amy-Louise and Joe Whittaker, “Evolution of terrorists’ use of the Internet,” Counterterror Business, 20 October 2017, <http://www.counterterrorbusiness.com/features/evolution-terrorists%E2%80%99-use-internet>.

⁵ Rita Katz, “TO CURB TERRORIST PROPAGANDA ONLINE, LOOK TO YOUTUBE. NO, REALLY.,” Wired (2018), https://www.wired.com/story/to-curb-terrorist-propaganda-online-look-to-youtube-no-really/amp?mbid=social_twitter_onsiteshare&_twitter_impression=true.

⁶ Kent Walker, “Four steps we’re taking today to fight terrorism online”, Google Blog (2017), <https://www.blog.google/around-the-globe/google-europe/four-steps-were-taking-today-fight-online-terror/>.

hate speak fuelling Myanmar's Rohingya Crisis.⁷ On the flip-side both Facebook and YouTube have been criticized for removing content posted by Rohingya activists highlighting the violence and persecution experienced by the Rohingya minority.⁸

4 Redirection: Demand-End Targeting

An alternative approach to disruption, known as the “redirect method”, has been piloted by Jigsaw, an initiative by Google in partnership with Moonshot CVE, Quantum Communications and a team of counter-narrative researchers.⁹ This approach identifies internet users that search for terms related to violent extremism and then targets them with adverts featuring counter-narrative content. As Jigsaw explains, “[t]he Redirect Method uses Adwords targeting tools and curated YouTube videos uploaded by people all around the world to confront online radicalization. It focuses on the slice of ISIS’ audience that is most susceptible to its messaging, and redirects them towards curated YouTube videos debunking ISIS recruiting themes.”¹⁰ Rather than making counter-narrative content themselves, the redirect method makes use of libraries of existing videos in Arabic and English. This is done for reasons of credibility, as a multinational organization is unlikely to be seen as a credible messenger itself. The results of an eight-weeks pilot showed that 320 000 individuals watched over half a million minutes of the 116 videos that had been selected for the adverts to refute ISIS’s recruiting themes.¹¹

The use of targeted ads by the re-direct method appears on the surface to elegantly sidestep many of the ethical dilemmas of disruption. As no content is removed, this avoids taking decisions based on the nature of the content. Similarly, as the counter-

⁷ Megan Specia, and Paul Mozur, “A War of Words Puts Facebook at the Center of Myanmar’s Rohingya Crisis”, *The New York Times* (2017), <https://www.nytimes.com/2017/10/27/world/asia/myanmar-government-facebook-rohingya.html>.

⁸ “Why are posts by Rohingya activists getting deleted?”, BBC News (2017), <https://www.bbc.com/news/blogs-trending-41364633>.

⁹ A similar pilot project has been set up by Microsoft on their search engine platform Bing in conjunction with the Institute of Strategic Dialogue (ISD).

¹⁰ “The Redirect Method: A blueprint for bypassing extremism,” The Redirect Method, 2, <https://redirectmethod.org/downloads/RedirectMethod-FullMethod-PDF>

¹¹ Ibid.

messages are placed as adverts, which an individual has to choose to click on and view, then watching the messages is an act of free will, and there is no “hidden hand” behind the scene deciding what you see.

At this stage, without more research, it is hard to know how effective the redirect method is. In terms of advertising metrics of number of views and click-through rates, it appears successful. However, this tells us little about what impact the videos had on those who viewed them, whether positive or negative¹². This echoes a wider problem of CT-CVE strategic communications, in which measurements and evaluations rarely go beyond engagement metrics, and typically fall short of measuring actual impact.

Ultimately, the redirect method and related approaches are essentially a targeted message campaign, exploiting keyword search technology of online search engines to target a specific audience. Whilst these approaches may satisfy “reach” criteria by connecting the message to the target audience, the construction of the message itself is important for maximising impact.

5 Counter-narratives: Hearts & Minds

The development of counter-narratives seeks to both address the supply side by confronting the messaging produced by violent extremists and the demand side by offering “shared” target audiences a rebuttal to violent extremist claims. However, there remains a lack of empirical evidence to support many of the underlying assumptions about counter-narratives and their impact.¹³ Andrew Glazzard summarises a key problem in the reasoning that underpins counternarrative development as the notion “that violent words lead to violent deeds, that counter-narratives can replace terrorist narratives, and that the actual threat of violent extremism can be

¹² Todd C. Helmus and Kurt Klein, *Assessing Outcomes of Online Campaigns Countering Violent Extremism: A Case Study of the Redirect Method*. Santa Monica, CA: RAND Corporation, 2018. https://www.rand.org/pubs/research_reports/RR2813.html.

¹³ Kate Ferguson, “Countering violent extremism through media and communication strategies: A Review of the Evidence” Research Associate, Partnership for Conflict, Crime and Security Research University of East Anglia (2016), <http://www.pacsresearch.org.uk/wp-content/uploads/2016/03/Countering-Violent-Extremism-Through-Media-and-Communication-Strategies-.pdf>

mitigated through discourse.”¹⁴ In practice, this means that as we do not fully understand the causal processes between the consumption of violent extremist material and engaging in violent extremism, we are working partially in the dark when trying to counter them.¹⁵

There is also too often a tendency for counter-narrative approaches to be based on the “hypodermic needle” model of communications. This approach takes a simplified linear perception of the information space, imagining the public as passive players, who all react similarly in changing their attitudes or behavior after being “injected/inoculated” with the same message.¹⁶ Finally, counter-narratives are inherently defensive in nature, in that they are focused on responding to the adversary’s narrative. In short, they “merely respond to the opposition’s message, allowing them to set the ground on which the communication battle will be fought and to maintain control of the narrative.”¹⁷ Although defensive messaging is an important part of any communication campaign, past communications campaigns have demonstrated that it is when offensive messaging is prioritized that a messaging effort is more likely to be effective.¹⁸

6 Rule #1: Do violent extremist propagandists no favours

It is important to begin by emphasising that almost any effort that disturbs the design, dissemination, reach and impact of violent extremist propaganda has some value that should never be flippantly dismissed. Any single approach – whether it be

¹⁴ Andrew Glazzard, “Losing the Plot: Narrative, Counter-Narrative and Violent Extremism”, *The International Centre for Counter-Terrorism – The Hague* 8, no. 8 (2017) <https://icct.nl/publication/losing-the-plot-narrative-counter-narrative-and-violent-extremism/>

¹⁵ Alastair Reed, “An Inconvenient Truth: Countering Terrorist Narratives – Fighting a Threat We Do Not Understand”, *ICCT Perspectives* (2018), <https://icct.nl/publication/an-inconvenient-truth-countering-terrorist-narratives-fighting-a-threat-we-do-not-understand/>

¹⁶ Cristina Archetti, “Terrorism, Communication and New Media Explaining Radicalization in the Digital Age”, *Perspectives on Terrorism*, Vol. 9, No. 1 (February 2015), pp. 49-59.

¹⁷ Alastair Reed, “IS Propaganda: Should We Counter the Narrative?”, *ICCT Perspectives* (2017), <https://icct.nl/publication/is-propaganda-should-we-counter-the-narrative/>

¹⁸ Haroro J. Ingram “A Brief History of Propaganda During Conflict: Lessons for Counter-Terrorism Strategic Communications”, *The International Centre for Counter-Terrorism – The Hague* 7, no. 6 (2016), <https://icct.nl/wp-content/uploads/2016/06/ICCT-Haroro-Ingram-Brief-History-Propaganda-June-2016-LATEST.pdf>

disruption, redirection, counternarratives or another strategy, including the one we propose here – will inevitably have effects that are relatively narrow and will need to be supplemented by other approaches. Three important questions need to be addressed when assessing the potential efficacy and efficiency of any approach to confronting violent extremist propaganda. First, what is the intent and purpose of that strategy? Second, how is that particular strategy being synchronized with other efforts? Third, and perhaps most importantly, what is the strategic logic underpinning that approach and what are its implications for practitioner mentality and team culture?

The trend towards disruption, redirection and counter-narratives, especially if deployed in unison, reflects a largely defensive posture and the application of a strategic logic of being responsive to the threat as it emerges and evolves.¹⁹ This posture and logic tends to fuel a mentality and culture amongst practitioners of “do no harm”. In combination, disruption, redirection and counter-narratives are a strong three-pillared defensive strategy. What we propose is for a more assertive posture that builds on these strong defensive foundations to frame the challenge of confronting violent extremist propaganda as a “competition” that requires both offensive and defensive strategies. So, while the combination of disruption, redirection and counter-narratives offers a strong defensive approach to confronting violent extremist propaganda, primacy should be given to offensive strategies that seek to proactively engage²⁰ with target audiences and prompt defensive counter-messaging from violent extremists. The strategic logic of this approach²¹ is based on a growing body of scholarship that understands a fundamental purpose of violent extremist propaganda to be the establishment of a “competitive system of meaning” (i.e. a prism through which perceptions are shaped) for its target audiences.²² Effectively and efficiently

¹⁹ Alastair Reed, Haroro Ingram and Joe Whittaker, “Countering Terrorist Narratives”, Policy Department for Citizens’ Rights and Constitutional Affairs European Parliament (2017), [http://www.europarl.europa.eu/Reg-Data/etudes/STUD/2017/596829/IPOL_STU\(2017\)596829_EN.pdf](http://www.europarl.europa.eu/Reg-Data/etudes/STUD/2017/596829/IPOL_STU(2017)596829_EN.pdf)

²⁰ Nafees Hamid, “Don’t Just Counter-Message; Counter-Engage”, *ICCT Perspectives* (2018), <https://icct.nl/publication/dont-just-counter-message-counter-engage/>

²¹ Haroro J. Ingram, “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, no. 6 (2017), <https://icct.nl/wp-content/uploads/2017/04/ICCT-Ingram-The-Strategic-Logic-of-the-Linkage-Based-Approach.pdf>

²² Haroro J. Ingram (2015) *The strategic logic of Islamic State information operations*, *Australian Journal of International Affairs*, 69:6, 729-75 <https://www.tandfonline.com/doi/full/10.1080/10357718.2015.1059799>; Alastair Reed and Jenifer Dowling, *The role of historical narratives in extremist propaganda*, *Defense Strategic Communications Journal* (2018)

“out-competing” violent extremist systems of meaning requires messaging and actions that not only undermine their credibility and veracity but promote an attractive competing system of meaning. We conclude this paper by outlining two strategies – the first related to campaign design, the second to message design – that are not only simple to deploy but help to encourage a more assertive posture, mentality and team culture.

7 The KISMI Approach to Campaign Roll-Out: Keep It Simple, Maximize Impact

One of the challenges when launching a strategic communications campaign, especially when the centrepiece of the effort is a long-form message like a written publication or video, is how to maximize its reach and impact. Drawing on empirical research from the social and behavioural sciences,²³ the KISMI approach to rolling-out a strategic communications campaign is to deploy messaging cycles dominated by PS-PS (Persuasive, Simple, Positive, Short) messages augmented by TANDEM (Thematically Accumulated, Narrative-Driven, Emotion Motivators) messaging. PS-PS messaging stands for:

- *Persuasive*: the message is deployed with the intent of leveraging rational- or identity-choice decision-making in its audiences. In other words, messaging is deployed with persuasive intent and never for simply informational purposes.
- *Simple*: the message is direct and unambiguous and does not rely on its audiences engaging in complex deliberative considerations in order to achieve its intent.

<https://www.stratcomcoe.org/alastair-reed-jennifer-dowling-role-historical-narratives-extremist-propaganda>; JM Berger, *Extremism*, MIT Press (2018) <https://mitpress.mit.edu/contributors/j-m-berger>.

²³ J.M. Berger, “Extremist Construction of Identity: How Escalating Demands for Legitimacy Shape and Define In-Group and Out-Group Dynamics”, *The International Centre for Counter-Terrorism – The Hague* 8, no. 7 (2017), <https://icct.nl/wp-content/uploads/2017/04/ICCT-Berger-Extremist-Construction-of-Identity-April-2017-2.pdf>; Haroro J. Ingram, “Deciphering the siren call of militant Islamist propaganda: Meaning, credibility & behavioural change”, *The International Centre for Counter-Terrorism – The Hague* 7, no. 9 (2016), <https://icct.nl/wp-content/uploads/2016/09/ICCT-Ingram-Deciphering-the-Siren-Call-of-Militant-Islamist-Propaganda-September2016.pdf>. Craig Whiteside, “Nine Bullets for the Traitors, One for the Enemy: The Slogans and Strategy behind the Islamic State’s Campaign to Defeat the Sunni Awakening”, *The International Centre for Counter-Terrorism – The Hague* 9 (2018).

- *Positive*: the message focuses on the benefits of certain rational- or identity-choice decisions. In contrast, negative messaging focuses on the detrimental impact of certain rational- or identity-choice decisions.
- *Short*: whether written, verbal or audiovisual, the message is brief in length.

TANDEM stands for:

- *Thematically Accumulated*: the message brings together the key themes which featured in the PS-PS messaging that preceded it in the messaging cycle. Put another way, the PS-PS messaging acts as a primer for the fusion of themes that appear in the TANDEM message.
- *Narrative-Driven*: the message ties together those aforementioned themes into a story. This narrative may use certain characters, issues or events to explore the relationship between those themes.
- *Emotion Motivators*: the message weaves emotion-based motivators pertinent to the target audience into its narrative.

Short, simple (PS-PS) messaging are deployed to prime target audiences for longer, complex (TANDEM) messages. This approach to campaign rollout is specifically designed for messaging to populations experiencing the cognitive impairment inherent to crisis.²⁴

For example, a capacity building project recently advised a civil society organisation to adopt the KISMI approach in its plans to hold an anti-IS-themed community event in a community that had been heavily impacted after IS forces had captured the city. The civil society organization engaged in a carefully planned short message campaign across offline and online mediums four weeks prior to the community event (i.e. the long-form message). This approach not only raised awareness about the event itself, maximizing turnout, but primed the potential audience for its key themes. After the event, the civil society organization rolled out a short-messaging campaign that ran for four to six weeks designed to reinforce the event's themes and transition into the next long-form message (video of the event).

²⁴ Haroro J. Ingram, "Islamic State's English-language magazines, 2014-2017: Trends & Implications for CT-CVE Strategic Communications", *The International Centre for Counter-Terrorism – The Hague* 8, no. 15 (2018), <https://icct.nl/wp-content/uploads/2018/03/ICCT-Ingram-Islamic-State-English-Language-Magazines-March2018.pdf>.

8 Leverage *Say-Do* Gaps

Given the “do no harm” mantra that often underpins CT-CVE strategic communications efforts, practitioners can find it difficult to identify messaging themes that can be deployed offensively and, thus, trigger counter-messaging from violent extremist adversaries. A relatively low-risk option is to identify examples where there is discord between violent extremist propaganda and their actions/reality on the ground. Leveraging say-do gaps is an opportunity to undermine the credibility of violent extremists as messengers and politico-military actors. Done effectively, it can also increase dissonance amongst sympathisers and supporters of violent extremists. To continue the example used previously, the capacity building project worked with civil society groups to design a short-messaging campaign after the community event which focused on highlighting IS say-do gaps. Local groups designed a range of messages that highlighted how both IS central and its local actors had lied to members, supporters and the broader population on issues ranging from theological and strategic to wages and the treatment of hostages.

9 Conclusion

This paper argued that the first rule of CT-CVE messaging should be to “do violent extremist propagandists no favors” as a means to drive a more assertive, competitively oriented posture for confronting violent extremist propaganda. Disruption, redirection and counter-narratives represent three important defensive strategies that, ideally, would be deployed in unison as the pillars of the defense component of a strategic communications effort. A more assertive posture requires complementary offensive efforts that, ideally, should dominate the campaign over time. We presented a campaign design and a message design strategy for informing a more holistic approach. The KISMI principle for campaign roll-out calls for the deployment of short, simple messaging as a means to prime target audiences for longer, more complex messaging. Leveraging violent extremist say-do gaps was identified as a message design strategy that can be a powerful way to erode the credibility of violent extremist actors and sow cognitive dissonance amongst sympathetic audiences. Ultimately, any single strategy in isolation will have its strengths and weaknesses. The

key to a more comprehensive approach to combating adversaries in the “information theatre” is to deploy a combination of approaches – both defensive and offensive with a priority on the latter – as the drivers of a more assertive posture inspired by the mantra of *do the adversary no favors*.