

Evaluating the Effectiveness of Online Counter-Messaging Campaigns Against Radicalisation and Violent Extremism

Valeria Rosato
University of Roma Tre

Abstract

This paper proposes a method for comprehensively evaluating the effectiveness of online counter- and alternative narratives (CN and AN, respectively) as part of countering violent extremism efforts. What are the main factors for an effective online counter-messaging campaign against radicalisation and violent extremism? Is it possible to measure the effectiveness of such online campaigns? The recent studies and research on online CN and AN agree in identifying a number of limitations in measuring abstract attitudinal or behavioural impacts. At the same time, however, they also recognize the enormous potential of such tools. The new effectiveness evaluation model proposed (the *GAMMA model revisited*) clearly shows the *realistic* objectives that an online campaign can set to contribute to the fight against violent extremism: 1. Raising awareness among the target audience (*Awareness*); 2. Encouraging online interaction and action (*Engagement*); 3. Promote the connection between online and offline actions (*Impact*).

Keywords: counter-narratives, violent extremism, social media, terrorism, online campaigns, radicalisation

Riassunto. *Valutazione dell'efficacia delle campagne online contro la radicalizzazione e l'estremismo violento*

Il presente articolo si propone di elaborare un metodo per la valutazione dell'efficacia delle contro-narrazioni online nell'ambito del contrasto all'estremismo violento. Quali sono i principali fattori che rendono efficaci le campagne di contro-narrazione online mirate al contrasto della radicalizzazione e dell'estremismo violento? È possibile misurare l'efficacia di tali campagne online? La maggior parte degli studi e delle ricerche sulla contro-narrazione online concordano nell'identificare una serie di limiti nella misurazione degli impatti di queste campagne sulle attitudini e sui comportamenti. Allo stesso tempo, però, riconoscono anche l'enorme potenziale di simili strumenti. Il modello di valutazione dell'efficacia che viene formulato (il *GAMMA model revisited*) identifica chiaramente gli obiettivi *realistici* che può perseguire una campagna online per contribuire al contrasto dell'estremismo violento: 1. Sensibilizzare il pubblico target (*Awareness*); 2. Incentivare l'azione e l'interazione online (*Engagement*); 3. Promuovere la connessione tra azioni online e offline (*Impact*).

Parole chiave: contro-narrative online, estremismo violento, radicalizzazione, social media, terrorismo

DOI: 10.32049/RTSA.2021.4.07

1. Introduction

Counter-messaging efforts continue to play a key role in contemporary policies to contrast terrorism and violent extremism. In response to the sophisticated communication skills shown by Islamic State (IS) terrorist group in online activities of recruitment and dissemination of digital contents, policymakers are paying increasing attention to the necessity to develop and augment online counter-messaging tools.

In asymmetric warfare such as in the case of the global war on jihadi terrorism, «violent

non-state political actors see messaging as crucial for winning the ‘heart and minds’ of supporters» (Ingram, 2016, p. 5) and for steering simple support towards violent collective action and insurgency. The creation of a «communicative action repertoire» (Bérubé and Dupont, 2019) and the exploitation of a wide array of channels of communication represent for instance a core component of the strategy of global jihadists movements, that have always invested substantial time and effort into propaganda.

In the years following the 11 September attacks, a new approach to terrorism emerged that focused on radicalisation understood as a linear process of “indoctrination” where vulnerable people in search of meaning would be manipulated by others intent upon radicalizing them (Silber and Bhatt, 2007). In order to be effective, the fight against terrorism had to aim at prevention, detection and management of the first signs of this process. Countering violent extremism or CVE along with preventing violent extremism (PVE) are terms that became commonplace over the past ten years (Williams, 2020). The terms relate to interventions that are aimed to prevent engagement in violent extremism (a term used in place of terrorism) but also to prevent recidivism for those already engaged in terrorist activity. CVE/PVE interventions could reasonably be said to be the mainstay of many State and regional (e.g., UN, EU) bodies’ responses to terrorism.

This model based on the concepts of “indoctrination” and “vulnerability” meant that policy and prevention programmes could focus on “counter-messaging”, which increasingly came to be understood as “counter-narratives” (McDonald, 2018).

Online campaigns and initiatives have thus proliferated in recent years, and these efforts can be divided into three categories based on the actors involved and the content disseminated (Briggs and Feve, 2013):

- 1) Counter-messaging (CN), aimed at deconstructing and discrediting violent extremist messages.
- 2) Alternative-messaging (AN), focused on the promotion of alternative narratives, positive values such as solidarity, and so on.
- 3) Government Strategic communications.

The goal of strategic communications is for the government to raise awareness of what it is doing and to build positive relationships with key constituencies. Counter-messaging, nevertheless, may also take the form of alternative narratives: positive stories about «social values, tolerance, openness, freedom and democracy» (Briggs and Feve, 2013). In other words, alternative narratives should be able to bridge the “us” versus “them” divide that is fostered by extremists (Schmid, 2014). Unlike the strategic communication (that is the purview of government), alternative narratives can be used by both government and civil society activists and groups.

However, the question as yet unanswered, is: how effective are these measures in preventing and combating radicalization and violent extremism? Evaluative research and analysis on Counter- and Alternative Narratives is limited and insufficient, both because this is a new field of study and because of the very nature of these initiatives. Indeed, if we consider that the objective of these efforts is to prevent and dissuade individuals from engaging in certain actions, any assessment method would need to measure a non-effect: non-radicalization. It is likewise difficult to assess the impact of counter-messaging with regard to the de-radicalization process, as this is a complex process entailing a large number of variables.

2. Methodology

This paper aims to introduce a model for a comprehensive evaluation of the effectiveness of online alternative and counter-narratives. To this end, we formulated two research questions:

RQ1: What are the main factors that make online counter-messaging campaigns against radicalisation and violent extremism effective?

RQ2: Is it possible to measure the effectiveness of such online campaigns?

To address these research questions, we have carried out the following steps:

1. Critical analysis of recent studies evaluating the effectiveness of online counter and

alternative narrative campaigns conducted as part of Counter Violent Extremism (CVE) efforts.

2. Identification and presentation of recent analytical models that may be useful in identifying the host of factors that play an important role in an online counter-messaging initiative.
3. Development of a guiding model for the implementation and evaluation of such campaigns.
4. Implementation of the proposed model by analysing several online awareness campaigns launched under the pilot study carried out as part of the European *Trivalent* project¹.

3. Critical analysis of the recent studies evaluating the effectiveness of online AN and CN campaigns

Radicalisation is a complex and gradual process in which an individual or group embraces radical ideologies or beliefs that endorse or use violence, including acts of terrorism, to achieve specific ideological or political goals. It is also necessary to point out that radicalism and terrorism are two concepts that cannot be equated. Radicalism typically precedes terrorism (Silber and Bhatt, 2007; Sageman, 2007) but there are also people with radical ideas and violent justifications who do not engage in violence or terrorism. There are many definitions of radicalisation and extremism in relation to terrorism, both at academic and institutional level (Borum, 2011). For simplicity and clarity, the following definitions of these two phenomena are proposed: extremism is a set of attitudes that lead people to embrace discourses and behaviours focused on a hate culture that fosters discrimination based on race, ethnicity, national origin, religious affiliation, sexual orientation, sex, gender, gender identity, serious disability or disease; radicalisation is a complex process of

¹ TRIVALENT «Terrorism Preventive via Radicalization Counter-Narrative», project funded by the European Union under the “Horizon 2020 Research and Innovation Programme”, grant agreement N. 740934 (<https://trivalent-project.eu>, 12/06/2020).

metanoia and socialisation that lead people to legitimize political violence as normal means of political action. Both of them are associated by the process of enemy-building and, within this standpoint, between the first and the second the difference is not qualitative but quantitative (Antonelli, 2019; 2021).

In the radicalization process an important role is played by narratives given their ability to offer a socially constructed version of reality that responds instrumentally to the propaganda goals of the narrator. According to Schmid (2014) a narrative is an easily-expressed story or explanation that organises people's experience and provides a framework for understanding events, while Braddock and Horgan (2016) define it as «any cohesive and coherent account of events with an identifiable beginning, middle, and end about characters engaged in actions that result in questions or conflicts for which answers or resolutions are provided».

Building on this finding, the focus of researchers and counter-terrorism practitioners in recent years has been on the potential of counter narratives in combating and preventing radicalization (Cherney, 2018; Cherney and Belton, 2019; Braddock, 2020). Generally speaking, CNs are narratives comprising content that contradicts the themes that makes up other narratives. In the context of violent extremism CNs represent attempts to directly, or indirectly, challenge violent extremist messages (Briggs and Fave, 2013) highlighting what is wrong with extremist ideologies and dismantle associated conspiracy theories (Schmid, 2014). These specific programs and campaigns can function both online and offline.

However, recent studies have shown little evidence of the effectiveness of counter narratives in reducing the risk of violent radicalization (Horgan and Braddock, 2010; Carthy *et al.*, 2020). The paucity of high-quality studies measuring these outcomes means that the evaluation cannot yet be considered conclusive and further research is needed.

Our focus in the present study is specifically on online counter-narrative with the goal of taking a state-of-the-art look at a still virgin and unexplored area of study. For this latter objective, recent studies focusing mainly on the analysis of the effectiveness of online campaigns to counter violent radicalization have been selected and critically analysed, highlighting their strengths and weaknesses.

In 2013, an early work by Reynolds and Tuck (2013) sought to offer guidance for monitoring and evaluating (M&E) a counter-narrative campaign. The core of CN campaign M&E lies in an analysis of the metrics from the websites and social platforms used. Social media metrics fall into two categories: *awareness* and *engagement*. The first refers to the total number of people who view the campaign contents (whether videos, ads, posts, websites or other digital material) and some information about those people; the second is defined as the volume and types of interaction between audience members, campaigners or campaign materials. Engagement can include everything from “likes” and “shares” to email responses, and can be either positive or negative. Knowing the number and nature of instances of engagement can help campaigners to understand their audience’s reactions to the campaign or its content. These metrics, which can be combined and analysed to form a comprehensive picture of a counter-narrative campaign’s performance, help campaigners to understand the *impact* they are having. Obviously, there is only so much that can be understood about the impact of such campaigns: analysing metrics can certainly offer insights into online impact, but it is very difficult to measure abstract attitudinal or behavioural outcomes. However, the M&E process can be improved and enriched by employing additional quantitative and qualitative research methods such as focus groups, in-depth interviews, online survey, geo-location analysis and sentiment analysis. Qualitative approaches in particular can provide a deeper understanding of how the public reacts to and perceives these campaigns.

More recently, empirical studies have been conducted on counter-narrative campaigns designed to prevent and fight terrorism on social media (Langer *et al.*, 2019; Silverman *et al.*, 2016; Speckhard, Shajkovci and Bodo, 2018; Speckhard *et al.*, 2018). In particular, Langer’s study of online campaigns showed which characteristics of social media posts produced the greatest reaction among the target group. 1. *Success*: posts in the form of reactions connected to a feeling of success; 2. *Determent*: personal fates such as stories of family members of the victims, former extremists etc. are popular. Indeed, the effectiveness of these campaigns’ narratives depends a great deal on the ability to provoke strong emotions. In terms of audience information, research shows the importance of *Self-*

promotion, an integrated call to action that offers interested people the opportunity to play an active role, while audiences respond less to hints and advice (*Suggestions*). Another important aspect for capturing attention and provoking greater reaction in all of these fields is visualization, i.e. the use of video and photographic material. In addressing the potential to improve impact on social media, the analysis sets off from the observation that these types of campaigns quite often have a low level of popularity. And finally, studies suggest that these initiatives must be integrated into other counterterrorism efforts, both by employing other effective ICTs (Information and Communication Technologies) and by intensifying real-world activities, especially interaction with potential target groups.

The serious limitations of this study lie in the fact that a large amount of data had to be excluded from further analysis due to the low rate of reaction on the part of the public. As most counter-narrative campaigns address a very specific target group and do not achieve high rates of response to their posts, however, it might be useful to examine and contrast the characteristics of messages with low response rates by applying a quantitative approach to a larger and more representative data sample.

Another research assessing the impact of online campaigns (Silverman *et al.*, 2016) presents the development, deployment and evaluation of three counter-narrative campaigns. Its initial research hypothesis is that, with even limited funding and a guide to support NGOs in launching counter-narrative projects on social media, appreciable results can be achieved in terms of *awareness*, *engagement* and *impact*. The analysis of social media metrics confirms high levels of *awareness* (the total number of people who view the campaign content) and *engagement* (volume and types of interaction). In particular, a qualitative analysis of the comments suggests that these campaigns display a capacity to promote critical reflection and the consideration of different points of view (“sustained engagements”). These campaigns also demonstrate that people undergoing processes of personal de-radicalization displayed a desire to contact the campaign-promoting organization to ask for assistance (*impact*). As for measuring the *impact* of counter-narratives, the authors themselves admit that assessment does not extend beyond impressions and engagement metrics. One positive outcome is undoubtedly the fact that it

succeeded in producing reactions, even negative ones, through comments, sustained engagement and conversations. The activation and promotion of online conversations continues to represent a useful tool that can potentially foster critical thinking through exposure to alternative views.

The International Center for the Study of Violent Extremism (ICSVE) has also conducted two highly interesting exploratory studies on efforts to combat ISIS ideology and its online dissemination. The first (Speckhard, Shajkovci and Bodo, 2018) presents the results of exploratory research on the effects of exposure to counter-narrative material on social media among ISIS group followers, promoters and sympathizers. The research not only explores the degree of radicalization and online behaviour of these individuals, but also examines the effects of posting counter-videos to their online pages, i.e. whether the proposed counter-narrative leads to any change in their observable behaviour. The research results indicate how this particular target (the same digital audience also addressed by terrorist groups) can be reached and demonstrate that people displaying radicalized online behaviour (through Facebook accounts) can easily be found and communicated with, spearheading counter-narrative materials. Obviously, little or nothing can be said about the long-term effects, in part because most of the individuals addressed in this study deleted the posted material and severed communication (by “unfriending” the sender or disabling their accounts). The best counter-narratives are those created by disillusioned insiders and not labelled as counter-narratives because these features grant them a better chance of reaching individuals who are already consuming terrorist materials. Moreover, the authors assert that this investigation, however exploratory, is nonetheless a useful starting point for future research, mainly based on the hypothesis that the dissemination of the true stories of ISIS deserters could produce effects among radicalized individuals.

The second ICSVE study (Speckhard *et al.*, 2018) focuses on countering ISIS ideology and its active recruitment in digital spaces by utilizing the same methods employed by the terrorist group. In this case, another exploratory research aims to identify the most suitable method for countering ISIS ideology on Facebook and ascertain if the counter-narratives used in these efforts might have an observable and measurable positive impact. This

exploratory study is important because it is one of the first online focus-groups of radicalized individuals conducted to test the exposure effects of CN materials and because it presents two significant results: it shows how this type of target audience can be reached and what kind of material (in particular, video clips of ISIS defector) is most effective for engaging this audience and their followers. Moreover, this study does not take into consideration the limit of not being able to evaluate long-term effects.

Another interesting study focuses on the effectiveness of counter-narratives by demonstrating the potential of combining online and offline tools. The study (Speckhard, Shajkovci and Ahmed, 2019) presents the results of a research project aimed at measuring attitudes towards extremist groups and the appeal of violent extremist ideologies such as ISIS among “vulnerable” Somali communities in the United States. It also makes a methodological argument for counter-narratives as an effective Counter Violent Extremism (CVE) tool. The focus group involved showing videos and visual materials with counter-narrative content to encourage discussion around issues of ISIS recruitment and ISIS-related radicalization. The authors point out some successful elements in the research: participation was active and satisfactory, and participants expressed gratitude for having been given the opportunity to express themselves in an environment they perceived as safe, given the particular sensitivity of the topics in question. Moreover, through dialogue and face-to-face interaction, the participants changed their positions and developed more solidly negative opinions about ISIS. According to the authors, this positive outcome depended in large part on the use of videos with emotionally evocative content and the conversational character of the experiment itself. Unfortunately, the research cannot be said to have intercepted with certainty a group of individuals “at risk” of radicalization. Despite its limitations, however, CN remains one of the main available tools for countering the appeal of violent extremism. The online dissemination of this type of material plays an important role in successfully reaching those who respond to such stimulus (in the exact same approach characterizing ISIS activity) with the aim of involving them further in online or offline discussions.

Finally, van Eerten and Doosje (2020) have carried out a significant examination of social media counter-messaging as a potential means of preventing or countering

radicalization in violent extremism. The study predominantly focuses on Salafi-Jihadi extremism countering efforts and examines the relative feasibility of preventing or countering radicalization and violent extremism via counter-messaging efforts. Given that radicalization is a very complex process involving many variables including possible exposure to online material, it follows that the task of assessing the effectiveness of counter-messaging in the framework of counter-narrative and prevention initiatives is likewise extremely complicated. Moreover, the greatest difficulty in this latter case lies in the very nature of the initiative: trying to measure a non-effect, i.e. the absence of radicalization. Along the same lines, however, the authors suggest the opposite assertion, that counter-messaging is ineffective, cannot be sustained either. They argue that counter-messaging has the potential to succeed only if it addresses people who show interest in and sympathy for extremist groups but are not yet involved in them, or possibly also already-radicalized individuals who show signs of doubt and reconsideration.

The study thus proposes several key principles for developing a social media counter-messaging campaign, beginning from the assumption that such a campaign is exclusively preventive in nature (although potentially including individuals who show signs of disengagement with the targeted campaigns). The study displays some limitations: the topic of narratives is addressed in various disciplines, making it difficult to collect up all the significant literature; it is a highly dynamic field characterized by constantly and rapidly updated data and results; there is a great deal of “grey literature” (not purely scientific) and limited resources for selecting experts and conducting interviews and focus groups with them; and, given the prevalence of literature focused on countering Jihadist propaganda, the results cannot be generalized to other extremist groups.

Considering the insights derived from our selected literature review, some strengths and weaknesses can be identified:

- the development and implementation of online CN and AN is still in its infancy, so the results of the research presented here must be treated with great caution. It is clear that the domain of counter-messaging is severely hampered by a lack of strong empirical research and thorough evaluation. These limitations are due in part to the difficulty of

securing radicalized individuals as research participants. Although the two exploratory studies carried out by the *International Center for the Study of Violent Extremism* (ICSVE) scholarly team (Speckhard *et al.*, 2018; Speckhard, Shajkovci and Bodo, 2018) demonstrate an ability to intercept radicalized online individuals, they also clearly show the difficulties in successfully involving them as active participants. As we have seen, the various theories analysed by van Eerten and Doosje (2020) clearly show that counter-messaging is a wholly ineffective tool in the de-radicalization process;

- CN and AN campaigns have proved useful as preventive tools. While it is clearly difficult to intervene in the thought or behaviour of radicalized individuals, there seems to be some promising room for manoeuvre in relation to the population of extremist group “fans” and sympathizers. In the case of Jihadist propaganda, these “Jihobbysts” (Brachman, 2009) are not directly affiliated with the groups in question but they consider themselves part of the movement and are active in disseminating materials online. Some of the studies presented here clearly show that these campaigns can offer alternative points of view, provoking doubt and encouraging the development of critical thinking. The evaluation of some CN and AN campaigns launched on social media (Silverman *et al.*, 2016), as we have seen, not only found high levels of *awareness* (the total number of people who view the campaign content) and *engagement* (volume and types of interaction), but also indicated that it is possible to foster critical reflection and the consideration of different points of view (“sustained engagements”). More interestingly, these campaigns also demonstrated that people undergoing a process of personal de-radicalization can be reached;
- the effectiveness of these tools in changing attitudes and behaviours remains difficult to assess. In general, it is extremely difficult to measure changes in attitudes and beliefs because such changes comprise a long-term process. It is even more challenging to measure and evaluate this aspect in the field of prevention in particular since the desired result (prevention, non-radicalization) is a non-event. The anonymity characterising the internet makes this task even more difficult: it is not easy to

determine precisely who is involved (or not) in a social media campaign. As van Eerten, J.J. and Doosje conclude in their study, it is virtually impossible to establish a direct connection between counter-messaging campaigns and the prevention of extremism (van Eerten and Doosje, 2020);

- the need for further testing and online and offline research. Much of the research presented here is focused on assessing the impact of online CN and AN campaigns and all admit to having limitations due to the very nature of the tools used. The main difficulty remains that of measuring changes in attitudes, beliefs and behaviours. In particular, scholars assert the importance of developing specific offline tools where such campaigns show a proven potential to intercept individuals who display some doubt about their chosen extremist group (Silverman *et al.*, 2016). As a case in point, the research carried out among Somali people in California, a community considered “vulnerable” to the influence of Jihadist propaganda, demonstrates the effectiveness of the focus group technique and face-to-face interaction in creating a comfortable and emotionally engaging environment among participants and, thanks in part to the presence of a moderator, stimulating intense conversation and the exchange of different points of view (Speckhard, Shajkovci and Ahmed, 2019);
- research confirms the usefulness of both quantitative and qualitative research tools in evaluating the effectiveness of campaigns. Finally, all the studies analysed here suggest that quantitative and qualitative research tools should be developed and integrated. *The Counter-Narrative Monitoring & Evaluation Handbook* (Reynolds and Tuck, 2016) clearly explains and details all the research tools that can effectively contribute to the monitoring and evaluation (M&E) process. Analysing the metrics can certainly offer insight into the online impact of CN and AN campaigns, but additional quantitative and qualitative research methods – such as focus groups, in-depth interviews, online surveys, geo-location analysis and sentiment analysis – are very important for enriching the evaluation itself. Obviously, each research tool has its pros and cons and must be selected and used according to the objectives of the research. For example, an online survey offers rapid results but cannot always reach the intended

audience, due in part to self-selection bias. Faced with the evident challenge of evaluating the effects of counter-messaging tools, one thing is clear: each research method can help to improve our knowledge.

The critical analysis of the selected literature clearly shows the difficulty and scarcity of research and models capable of measuring the effectiveness of the counter- and alternative narrative efforts carried out to date in tackling violent extremism. Before asking if the effectiveness of such online campaigns can be measured, however, it is appropriate to appraise the main factors for an effective online counter-messaging campaign against radicalization and violent extremism.

4. Online counter-messaging: the relationship between effectiveness and credibility

Recent studies agree in clearly showing that the role of *credibility* and *trust* in the entire online communication process is central and must be analysed in more depth. We think that currently, this aspect represents the real challenge in potentially developing more effective counter-messaging strategies against radicalisation processes.

A recent study by Braddock and Morrison (2018) focuses on the link between credibility and effectiveness in online counter-narratives against terrorism, showing that the transformative power of such messages depends not only on their content but also on how the target audience perceives the source. The study analyses and highlights the failure of some counter-narrative programs due to a perceived lack of source credibility and reliability on the part of the selected target². The messages in these programs may be extremely accurate, but audience's lack of confidence gives rise to a rejection response, thus nullifying any persuasive effect.

The authors thus develop useful guidelines for constructing credible and effective

² For instance the case of the initiative *Think Again Turn Away*, a counter-narrative social media program of the US State Department's Center for Strategic Counter-terrorism Communication established in late 2013 and operating as a Twitter account (https://twitter.com/ThinkAgain_DOS, 12/10/2020) that posted counter-messaging content aimed at ISIS. This program did not appear to achieve any clear results and it was criticized for its poor production quality in comparison to ISIS propaganda product as well as for its counterproductive effects.

communication strategies, beginning from and linking two interesting models: the *Arc of Terrorism* model developed by Horgan (2014) and the *MAIN* model proposed by Sundar (2008). Horgan's model focuses on people's involvement in and disengagement from terrorism, demonstrating that terrorists cultivate confidence in their propaganda to improve effectiveness and that such confidence in the content sources is fundamental for disengagement as well. Sundar's model addresses the credibility of online sources by identifying the specific content in and structural characteristics of online messages that affect the degree of credibility users attribute to message sources. In exploring the credibility of the internet, this study identifies 4 specific types of affordances (modality, agency, interactivity and navigability) that can have significant psychological effects on online message users. That is, these four structural aspects of internet-based communication technologies can be very useful for understanding the online contexts of terrorist communication and counter-narrative campaigns in that they may influence users' perceptions of source credibility.

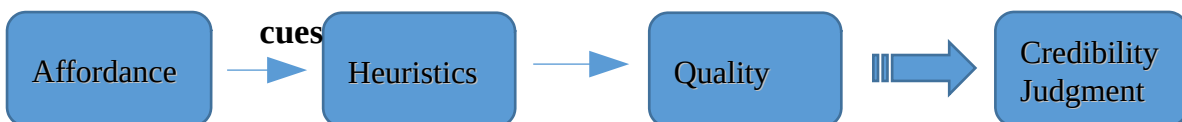


Fig. 1: Overview of the MAIN model. Source: Braddock and Morrison, 2018.

According to Sundar, therefore, psychological and behavioural effects vary depending on how these structural aspects occur in the medium itself. In particular, the affordances of “modality” and “agency” are relevant for the content of counter- and alternative narratives, while the aspects of “interactivity” and “navigability” refer to the way CN and AN are integrated into online interfaces.

Credibility and trust are thus fundamental variables in the communication process and can only be achieved by taking into account the close interrelation between socio-technical systems and narratives. This consideration is also at the core of the new narrative ecosystem paradigm developed by a Università Cattolica del Sacro Cuore research group (Lombardi, Lucini and Maiolino, 2020) as part of a pilot study launched under the framework of the

TRIVALENT (*Terrorism Preventive via Radicalization Counter-Narrative*) project and involving 300 students enrolled in a B.A. course. Specifically, a format model was developed as part of the innovative paradigm of the narrative ecosystem that may be useful in understanding the communication processes that influence both processes of radicalization and extremist communication and proactive, adaptive programmes of combatting and prevention.

The format is composed of two fundamental elements:

- 1) the communication ecosystem;
- 2) the narrative-scope.

The first refers to the context in which narratives operate, interacting with complex and interconnected, mutually influencing socio-technical systems. Moreover, the authors argue, the social context in which people live is characterized by a specific form of vulnerability. The last important element they identify is the pulverization of the audience, thus suggesting it is necessary to focus on micro-targeting practices and selected audience groups. As detailed below in the analysis of the selected campaigns developed in the pilot study, therefore, violence is conceptualized as multidimensional and dependent on personal, social and contextual variables. In particular, *emotions*, *feelings* and *affects* play a fundamental role in defining the characteristics of the target audience and the communication ecosystem. According to the F.E.A.R. model developed by this same research group, feelings (personal), emotions (social) and affects (pre-personal) are key in analysing violent behaviours.

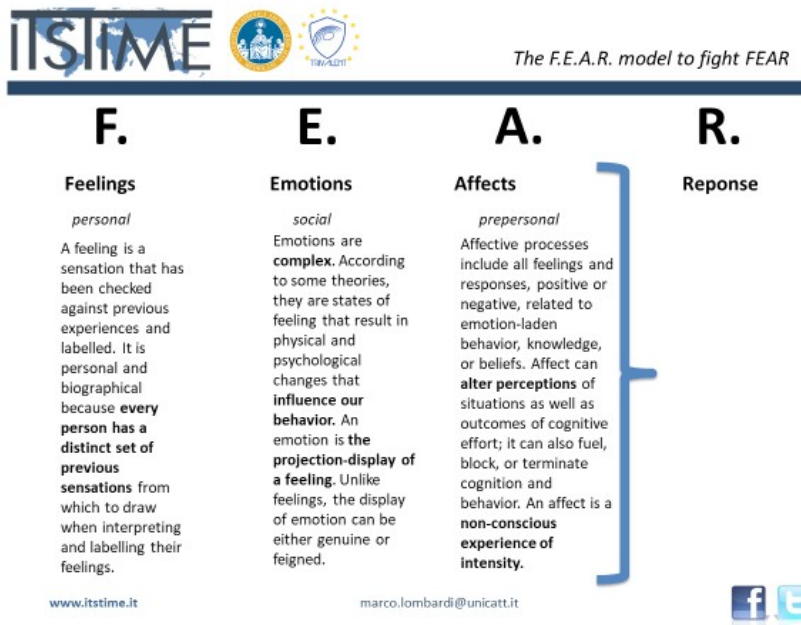


Fig. 2: The F.E.A.R. model to fight FEAR. Source: Lombardi, Lucini and Maiolino, 2020

The format model corresponds to the “R” of F.E.A.R. and has two fundamental characteristics: it is context-sensitive and target-oriented. This model is useful for guiding the communication processes that influence both the processes of radicalization and extremist communication, and the actions of contrast and prevention. The format model is thus composed of two interrelated components to be adapted to different audiences, scenarios and socio-technical systems, graphically summarized as follows:

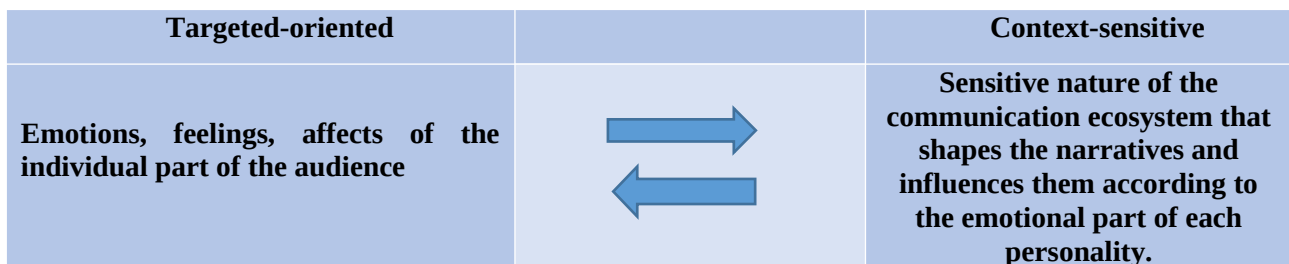


Figure 3. Format model

5. Evaluating online counter-messaging: the “GAMMA model revisited”

On the basis of our review of existing literature on the subject and the results of the pilot study carried out as part of the *Trivalent* project, we present our evaluation and measurement proposal broken down by key element along the lines of the GAMMA+ model formulated by the European Commission’s *Radicalisation Awareness Network* (RAN)³: Goal, Audience, Message, Messenger, Media, and Action, plus Monitoring and Evaluation (RAN 2017, 2019). This revisited model is proposed as a guideline for developing and implementing, monitoring and evaluating an effective alternative and counter-narrative campaign. For each element we briefly outline our proposal for how to measure the effectiveness of an online campaign.

Goal. According to RAN guidelines, effective communication campaigns must have clear, realistic and measurable objectives. Unfortunately, all the studies and research carried out so far on the effectiveness of CNs and ANs have clearly displayed serious limitations as to the possibility of measuring the effectiveness of such campaigns, in particular online campaigns, in any way other than partially. In reality such campaigns can realistically achieve some objectives, such as that of measuring levels of awareness and engagement through social media metrics, but they can hardly measure impact understood as the ability to modify the minds and behaviours of the individuals involved. As demonstrated by some pilot studies we have analysed, measurement of this kind might become more feasible by shifting the focus from online to offline actions. Given the development of technology, online initiatives have enormous potential in terms of awareness-raising and intercepting target groups; beginning by involving a small target group, offline initiatives can instead make use of a series of qualitative survey techniques designed to assess and measure possible changes in individuals' attitudes and behaviours.

Therefore, our analysis suggests that an online campaign may set itself three main realistic and achievable objectives:

³ The Radicalization Awareness Network (RAN) is a network supported and funded by the European Commission (https://ec.europa.eu/home-affairs/what-we-do/networks/radicalisation_awareness_network_en, 12/10/2020). Its purpose is to network practitioners from around Europe working on the prevention of radicalization.

1. Raising awareness among the target audience
2. Encouraging online interaction and action
3. Promoting the connection between online and offline actions

As we will see, to achieve these objectives the common thread linking all the key elements of a campaign against violent extremism is confidence and credibility. To be effective, a communication campaign must be perceived as credible and reliable in terms of both its content and the media it employs.

Audience. All CN and AN implementation guidelines agree on the importance of identifying a target audience to ensure the effectiveness of the campaign. As we have seen, a communicative context characterized by greater audience heterogeneity and pulverization is more vulnerable to misinformation and is also more difficult to control. Therefore, it is useful to identify communities, groups, and specific categories to which our message is addressed, i.e. focusing on microtargeting practices and segments of selected audience groups. As demonstrated by the pilot study, the concept of violence is multidimensional and directly dependent on multiple variables, personal, social and contextual, that must be taken into account when designing an online communication campaign.

The studies agree in recognizing that young people are an especially relevant category by virtue of their vulnerability to the phenomenon of violent extremism. In addition, the younger generation has deep familiarity with technology and can more easily be reached and involved in online communication campaigns. On the other hand, however, they are a particularly vulnerable segment of the population due to their as-yet-immature cognitive abilities and, therefore, more subject to the risks inherent in using new technologies (see *Media*). As we will see below, this last aspect is very important when the campaign is based on peer-to-peer communication (see *Messenger*).

Message. As many studies have shown, the main alternative and counter-narrative strategies employed to date have a number of limitations and problematic aspects. Counter-messaging, aimed at deconstructing and discrediting violent extremist messages, has displayed a series of boomerang effects: first, the risk of further polarizing the social context and augmenting the strength and appeal of extremist messages. Due to its oppositional

nature this approach may potentially repeat and reinforce the very messages it seeks to counteract, and there is no question it has been found completely ineffective on individuals who are already radicalized (van Eerten and Doosje, 2020). Alternative-messaging, focused instead on promoting narratives featuring positive values such as solidarity and so on, has limited effectiveness in fostering audience involvement. This limit stems from the fact that negative messages tend to resonate more than positive ones in online communication. Finally, government strategic communication suffers from a considerable lack of confidence and credibility due to the very nature of the source, as governments may be perceived as distant from the target audience and not impartial.

As illustrated by the new narrative ecosystem paradigm and Format Model, communications aimed at conditioning people's actions to combat and prevent extremism must take into account the personal, social and contextual variables that characterize the target audience and specific type of communication ecosystem. An effective alternative communication strategy would thus go beyond predefined categories such as today's CNs, ANs and government strategic communication, but it should be calibrated to the sensitivities and needs of the target audience.

What all the studies examined show and the F.E.A.R. model and proposed Format model confirm is that content and messages with a strong emotional impact have enormous potential. Extremist narratives themselves clearly illustrate that emotional logics are hugely effective because they respond to certain basic human needs (a demand for meaning, adventure, perceived deprivation, and camaraderie, to name just a few). Therefore, persuasion is facilitated by the immersion that can be achieved using vivid language, visual materials and the involvement of individuals and personalities similar to the target audience.

Messenger. As noted when analysing the previous elements, it is important that actors who can transmit confidence and credibility to the target audience be used as campaign sources. In particular, we have seen that individuals perceive a similar audience, and therefore similar sources, as more credible. Young people are the target audience of most of the counter-messaging efforts; it follows that peer-to-peer communication can have considerable potential, but also entail considerable risk given this category's vulnerability

due to their age and more limited ability to consciously manage the potential dangers inherent in using new technologies. Moreover, the experience and authority transmitted by the source play an important role. In the case of alternative and counter-narratives, messages from leaders, former terrorist group members, respected personalities from inside the online community, so-called “influencers”, etc., could exercise greater dissuasive power.

An important distinction can be made between formal and informal counter-messaging in terms of sources and credibility. Communication campaigns against violent extremism suffer a significant decrease in public confidence especially when they are launched or sponsored by governments. Such narratives risk being equated with propaganda and rendering attempts to counter extremist ideologies counter-productive. Civil society suffers less from this lack of credibility, although the campaigns of large organisations are often perceived as lacking in objectivity and reliability because they are aligned with or financially supported by governments. One field that has received less attention, however, is informal counter-messaging (Lee, 2019), that is, the array of initiatives in this domain carried out by individual citizens without any connection to governments or larger organisations. This is a vast, informal sector of actors producing digital content that is critical of extremist messages. Much of this informal material is disseminated on social media, and it is very successful. Informal actors are perceived as “free”, authentic and, therefore, more in line with their audiences. Moreover, studies on online persuasive communication also show that perceived confidence (among other aspects) is reinforced if the source is presented through dynamism, likeability and goodwill (Wathen and Burkell, 2002). Informal counter-messaging is often based on parodies and imitations that become “viral”, reaching a very wide audience and, often unintentionally, creating a “natural” counterweight to the spread of extremist messages. Informal counter-messaging also involves limits and risks, but its power has yet to be explored and analysed through further research, especially in view of possible synergy and collaboration with counter-messaging organizations.

Media. A wide range of media can be employed to convey narratives against violent extremism. Our analysis focuses on online communication and identifies social media as the

most effective tool thanks to certain of its specific traits: diffusion, versatility and, above all, interactivity. The latter aspect is fundamental if we consider the importance of interacting with and fostering emotional involvement among the public 2.0 platforms such as Facebook, Twitter or Instagram that facilitate person-to-person interaction promote perceptions of authority and reliability among users and can thus enhance the credibility of the campaign itself. In addition, however, it is important to acknowledge the risks inherent in new technologies, such as the “algorithmic dimension”. Social media algorithms are based on a selection mechanism that tends to offer personalized content, thus constantly exposing individuals with certain opinions to content in line with their beliefs. In so doing they limit the spectrum of different opinions, fostering polarization and prejudice. This process of “personalization” is even truer for platforms such as Youtube that are widely used by extremist actors, young people, and CN and AN campaign designers. According to the selective exposure paradigm, therefore, automatic algorithms offer users videos and information that match their interests and orientations, but it is also necessary to consider that the online content being suggested depends not only on individual selection but also the use of keywords. A study on Youtube’s role in counter-messaging (Schmitt *et al.*, 2018) clearly shows that inserting common keywords such as “ISIS” generates results including both extremist content and counter-messaging. The fundamental issue raised by the research is that the counter-messaging material on the web is clearly inferior to extremist propaganda, creating a marked imbalance. Paradoxically, then, there is a risk that counter-messaging campaigns might expose users to counter-productive messages and content, thus increasing the chances of promoting extremist ideas. The key point to consider is thus being aware of these dynamics, knowing how social media work and, obviously, developing a critical understanding of the manipulative capabilities of the messages to which the user is exposed.

Action. The last element of a counter-messaging initiative is action, understood as the expected response on the part of users. In the case of an online campaign, actions can be divided by level of involvement and range from a minimum level of involvement to outright action. Campaigns can thus focus exclusively on raising general awareness, set objectives

such as obtaining likes and shares for their contents, or aim at active involvement through more in-depth interactions and online discussions, organised encounters and the connection with offline actions and initiatives. An important objective, therefore, is to foster proactive involvement that stimulates users (but also the producers and consumers of online campaigns) to become involved.

Monitoring and Evaluation. Naturally, the *Monitoring & Evaluation* phase anticipates, accompanies and concludes each communication campaign. Effectiveness evaluation and measurement obviously depends on the specific objectives the producer has set, thus bringing to a close the circle of the GAMMA model.

Our analysis clearly shows that realistic objectives for an online counter-messaging campaign would be to raise awareness and encourage action, both online and offline. The effectiveness of these initiatives can be measured in part through social media metrics that aid in estimating the level of awareness and involvement, but nothing can really tell us about changes in target audience attitudes and behaviours. The challenge is to design a campaign aiming to be as *credible* and *effective* as possible thanks to the key elements highlighted by the analysis of research and studies conducted so far.

5.1 Analysing P2P communication campaigns according to the “revised GAMMA model”

To apply the “revised GAMMA model”, we analyse some online campaigns developed as part of the pilot study that led to the development of the new narrative paradigm ecosystem mentioned above.

One important result of the P2P Campaign is that it shows how violence and extremism are depicted among the students and what outputs they developed. The project generated 43 peer-to-peer communication campaigns focusing on the following macro thematic categories:

1. Bodily Physical Violence;
2. Hate Speech;

3. Violence against disabled people;
4. Solidarity and inclusion;
5. Insults;
6. Psychological violence;
7. The environmental emergency;
8. Violence against artistic and cultural heritage.

These eight categories illustrate the variety of forms of violence the students identified, rendering evident what constitutes extremism and harm for them. In developing these definitions, they have been oriented by the following factors:

- personal and biographical experience with the types of violence;
- past experiences, first-hand or indirect, with a specific type of violence;
- personal features;
- perceived risk of exposure;
- cultural ways of defining what does and does not constitute violence and extremist behaviour;
- the communication strategies through which a given violent message and counter-communication are disseminated.

For our analysis, we selected the 5 most significant campaigns⁴:

1. *What about you*⁵, a campaign that aims to raise awareness, especially among young people, about the phenomenon of victim blaming;
2. *Let's talk about that*⁶ is aimed at young women and aims to raise awareness about violence against women;
3. *Haters fighters*⁷ is a project to raise awareness and combat cyberbullying;
4. *Cattivissimo me*⁸ is a campaign to report and raise awareness about violence by law

⁴ All of these campaigns have been launched on the Facebook page “LIMED-Storie Responsabili”, a public communications page maintained by the Theories and Techniques of Media Communication class at the Catholic University of Milan, “LIMED-Storie Responsabili” (https://www.facebook.com/StorieResponsabili/?ref=py_c, 12/10/2020).

⁵ <https://www.instagram.com/whataboutyou>, 12/10/2020.

⁶ https://www.instagram.com/letstalkaboutthat_/, 12/10/2020.

⁷ https://www.instagram.com/haters_fighters_/, 12/10/2020.

⁸ https://www.instagram.com/cattivissimome_unicatt/ and <https://m.facebook.com/cattivissimome2020>, 12/10/2020.

enforcement agencies in prisons and barracks;

5. *Burn the violence*⁹ is an advertising campaign aimed at raising awareness about online violence, targeting young people and adults.

All of these campaigns seek to raise awareness and encourage action to address different types of violence, in particular victim blaming, violence against women, cyberbullying, prison violence and online violence. The heterogeneity of content, that is, the different types of violence identified to develop the awareness-raising campaigns, clearly shows that the campaign designers choices depended on a number of personal and contextual variables underlying their very definition of the concept of violence. The initiatives are based on peer-to-peer communication in which young people act as the sources and producers of the communication campaigns and are also identified as the main target audience.

The communication channel privileged in this case is the social network *Instagram*, chosen for its accessibility, strong visual impact, interactivity and widespread use by young people. In addition, this online platform allows users to attach other communication formats and to promote and link to initiatives on other channels and media. This latter capacity is evident if we consider the opportunities to advertise and refer users to content on other media formats, such as television, that are much more popular among the target audience and deal with the same topics. For instance, the *Let's talk about that* campaign references a television series (*You*) that explores the risk of using new communication technologies and falling victim to stalkers; it also refers, via hashtag, to the page of a well-known television series that is very popular with young people.

As regards the communicative style, these campaigns favour short images and videos that convey clear and straightforward messages, informational content created by professionals and activists, and emotional content conveyed through images and the first-hand accounts of victims.

Finally, user response is in line with the campaign's objective of raising awareness and can be measured through engagement (likes, shares, views and comments responding to the campaign messages).

⁹ https://www.youtube.com/watch?v=RNeX_OjR-Ko, 12/10/2020.

NAME	Goal	Audience	Message	Messenger	Media	Action
<i>What about you</i>	Promoting awareness about the phenomenon; dismantling prejudice	Both young people and young adults	Victim blaming	Young students	Instagram Flier	Likes, shares, comments and online discussion
<i>Let's talk about that</i>	Promoting awareness about the phenomenon; dismantling prejudice	Both young people and adults from 16 to 35 years old, with specific attention to the women	Violence against women	Young students	Instagram #LTAT	Likes, shares, comments and online discussion
<i>Haters Fighters</i>	Promoting awareness about the phenomenon; dismantling prejudice; tackling the phenomenon	Both young people and adults	Cyberbullying	Young students	Instagram #StopCyberBullying	Likes, shares, comments and online discussion
<i>Cattivissimo me</i>	Promoting awareness about the phenomenon; dismantling prejudice; tackling the phenomenon	Both young people and adults	Violence in prison	Young students	Instagram Facebook	Likes, shares, comments and online discussion
<i>Burn the violence</i>	Promoting awareness about the phenomenon; dismantling prejudice; tackling the phenomenon	Both young people and adults	Flaming	Young students	YouTube	Visualization Likes shares

Tab. 1 Analysis of P2P campaigns according to the “revised GAMMA model”

6. Conclusions

All of the studies and research on online CN and AN analysed here identify a number of limitations on the prospects of measuring abstract attitudinal or behavioural impacts. At the same time, they also recognize the enormous potential of these tools in terms of preventive action.

If well-designed, awareness-raising campaigns against radicalization and violent extremism can offer alternative point of views, instilling doubt and encouraging the development of critical thinking.

Social media metrics are a useful instrument for quantitatively measuring the effectiveness of an online campaign in terms of *awareness* and *engagement*, but they rarely allow us to assess *impact* in terms of changing beliefs, attitudes and behaviours. The latter type of evaluation can only be achieved through qualitative analysis methods such as focus groups, in-depth interviews etc., and the development of specific offline tools.

The “revised GAMMA model” clearly indicates the *realistic* objectives that an online campaign can set for itself so as to effectively contribute to the fight against violent extremism:

1. Raising awareness among the target audience (*Awareness*).
2. Encouraging online interaction and action (*Engagement*).
3. Promoting the connection between online and offline actions (*Impact*).

In relation to these objectives, each campaign must choose and modulate the key elements comprising a communication campaign (Audience, Message, Messenger and Media) in such a way as to ensure significant levels of *credibility* and *reliability*. The analysis of the case study, for example, clearly shows that peer-to-peer communication has considerable potential in this area if calibrated to the characteristics of the target audience and the communication context. In addition, the proposed Format Model should be taken into consideration as it represents an interesting new element in efforts to counter extremist communication.

Here below we summarize tools for measuring effectiveness based on the objectives of

an online campaign:

GOALS	ACTIONS	TOOLS
<p><i>Awareness</i></p> <p>Number of individuals who have been reached</p>	Impressions, reach and views	Social media and website analytics services
<p><i>Engagement</i></p> <p>Number of individual interactions</p>	Likes, comments, shares and retweets/reposts	Social media and website analytics services
<p><i>Impact</i></p> <p>Connection to online and offline actions</p>	Comments, interactions, and discussions	Quantitative and qualitative data analysis: geo-location analysis; sentiment analysis, online surveys, in-depth interviews and focus groups

Tab. 2 Tools for measuring the effectiveness of an online campaign

In conclusion, the state-of-the-art literature on CN/AN clearly shows the persistent difficulty in building models capable of assessing and measuring the effectiveness of online campaigns against violent extremism and radicalisation. The main finding of the analysis is that at present the real challenge in countering the negative effects of extremist messages lies in combining and linking online and offline PVE/CVE efforts. These efforts should be directed primarily at strengthening people's social cognitive resilience to violent extremism, especially among young people, by fostering a greater critical understanding of the manipulative nature of messages and a greater awareness not only of the role of the internet and its channels, but also of how it works.

Reference List

Antonelli F., ed. (2019). *Working Papers in Terrorism Studies: The Present And The Future Of Violent Radicalization In Europe*. Roma: Roma Tre Press.

- Antonelli F. (2020). *Radicalizzazione*. Milano: Mondadori.
- Bérubé M., Dupont B. (2019). *Mujahideen Mobilization: Examining the Evolution of the Global Jihadist Movement's Communicative Action Repertoire*. *Studies in Conflict & Terrorism*, 42, 1-2: 5. DOI: 10.1080/1057610X.2018.1513689.
- Borum R. (2012). *Radicalization into Violent Extremism I: A Review of Social Science Theories*. *Journal of Strategic Security*, 4, 4: 7. DOI: 10.5038/1944-0472.4.4.1.
- Brachman J.M. (2009). *Global jihadism: Theory and practice*. London: Routledge.
- Braddock K. (2020). *Weaponized Words: The Strategic Role of Persuasion in Violent Radicalization and Counter-Radicalization*. Cambridge: Cambridge University Press.
- Braddock K., Horgan J. (2016). *Towards a Guide for Constructing and Disseminating Counternarratives to Reduce Support for Terrorism*. *Studies in Conflict & Terrorism*, 39, 5: 381. DOI: 10.1080/1057610X.2015.1116277.
- Braddock K., Morrison F.J. (2018). *Cultivating Trust and Perceptions of Source Credibility in Online Counternarratives Intended to Reduce Support for Terrorism*. *Studies in Conflict and Terrorism*, 43, 6: 468. DOI: 10.1080/1057610X.2018.1452728.
- Carthy S.L., Doody C.B., Cox K., O'Hora D., Sarma K.M. (2020). *Counter-narratives for the prevention of violent radicalisation: A systematic review of targeted interventions*. *Campbell Systematic Reviews*, 16, 3. DOI: 10.1002/cl2.1106.
- Cherney A. (2018). *Evaluating Interventions to Disengage Extremist Offenders: A Study of the Proactive Integrated Support Model (PRISM)*. *Behavioral Sciences of Terrorism and Political Aggression*, 12, 1: 17. DOI: 10.1080/19434472.2018.1495661.
- Cherney A., Belton E. (2019). *Evaluating Case-Managed Approaches to Counter Radicalization and Violent Extremism: An Example of the Proactive Integrated Support Model (PRISM) Intervention*. *Studies in Conflict and Terrorism*, 44, 8: 625. DOI: 10.1080/1057610X.2019.1577016.
- Horgan J. (2014). *The Psychology of Terrorism*. London: Routledge.
- Horgan J., Braddock K. (2010). *Rehabilitating the Terrorists?: Challenges in Assessing the Effectiveness of De-radicalization Programs*. *Terrorism and Political Violence*, 22, 2: 267. DOI: 10.1080/09546551003594748.

- Ingram H.J. (2016). *Deciphering the Siren Call of Militant Islamist Propaganda*. International Centre for Counterterrorism Research Paper. The Hague: ICCT. DOI: 10.19165/2016.1.12.
- Langer A., Kaufhold M.A., Runft E.M., Reuter C., Grinko M., Pipek V. (2019). Counter Narratives in Social Media - An Empirical Study on Combat and Prevention of Terrorism. In Franco Z., González J.J., Canós J.H., eds., *Social Media in Crises and Conflicts*. Proceedings of the 16th ISCRAM Conference. València: ISCRAM. Retrieved from: http://www.peasec.de/paper/2019/2019_Langeretal_Counternarratives_ISCRAM.pdf (22/12/2021).
- Lee B.J. (2019). Informal Countermessaging: The Potential and Perils of Informal Online Countermessaging. *Studies in Conflict & Terrorism*, 42, 1-2: 161. DOI: 10.1080/1057610X.2018.1513697.
- Lombardi M., Lucini B., Maiolino M. (2020). Beyond counter- and alternative narratives to tackle extremism: the new Format model. *Sicurezza, Terrorismo e Società*, 11, 1: 7. Retrieved from: http://www.sicurezzaeterrorismosocieta.it/wp-content/uploads/2020/06/eBookInterno_SicTerSoc11_Beyond-counter-and-alternative-narratives.pdf (10/09/2020).
- McDonald K. (2018). *Radicalization*. Cambridge: Polity Press.
- Radicalisation Awareness Network - RAN. (2017). *How to measure the impact of your online counter - or alternative narrative campaign*. Ex Post Paper, RAN C&N meeting, 13 and 14 February, Brussels. Retrieved from: https://ec.europa.eu/home-affairs/sites/homeaffairs/files/docs/pages/201702_ran_how_to_measure_impact_of_online_campaign_en.pdf (03/10/2020).
- Radicalisation Awareness Network - RAN. (2017). *RAN Guidelines for effective alternative and counter-narrative campaigns (GAMMMMA+)*. RAN Issue Paper, 31/12. Retrieved from: https://ec.europa.eu/home-affairs/system/files/2020-09/ran_cn_guidelines_effective_alternative_counter_narrative_campaigns_31_12_2017_en.pdf (12/09/2020).
- Radicalisation Awareness Network - RAN. (2019). *Monitoring & Evaluating counter- and alternative narrative campaigns*. Retrieved from: <https://ec.europa.eu/home-affairs/sites/homeaffairs/files/what-wedo/networks/>

radicalisation_awareness_network/about-ran/ran-c-andn/docs/

ran_cn_ex_post_evaluating_campaigns_berlin_210219_22_en.pdf (10/10/2020).

Reynolds L., Tuck H. (2016). *The counter-narrative monitoring & evaluation handbook*. London: Institute for Strategic Dialogue (ISD). Retrieved from: <https://www.isdglobal.org/wp-content/uploads/2017/06/CN-Monitoring-and-Evaluation-Handbook.pdf> (22/12/2021).

Sageman M. (2004). *Understanding Terror Networks*. Philadelphia: University of Pennsylvania Press.

Schmid A.P. (2014). *Al-Qaeda's "Single Narrative" and Attempts to Develop Counter-Narratives: The State of Knowledge*. International Centre for Counterterrorism Research Paper. The Hague: ICCT. Retrieved from: <https://www.icct.nl/download/file/AP-Schmid-Al-Qaedas-Single-Narrative-January-2014.pdf> (10/01/2021).

Schmitt J.B., Rieger D., Rutkowski O., Ernst J. (2018). Counter-messages as Prevention or Promotion of Extremism?! The Potential Role of YouTube: Recommendation Algorithms. *Journal of Communication*, 68, 4: 780. DOI: 10.1093/joc/jqy029.

Silber M.D., Bhatt A. (2007). *Radicalisation in the West: The Homegrown Threat*. New York: New York City Police Department. Retrieved from: <https://info.publicintelligence.net/NYPDradicalization.pdf> (22/12/2021).

Silverman T., Stewart C.J., Amanullah Z., Birdwell J. (2016). *The impact of counter-narratives. Insights from a year-long cross-platform pilot study of counter-narrative curation, targeting, evaluation and impact*. London: Institute for Strategic Dialogue. Retrieved from: https://capve.org/components/com_jshopping/files/demo_products/Impact-of-Counter-Narratives_ONLINE_1.pdf (22/12/2021).

Speckhard A., Shajkovci A., Bodo L. (2018). *Fighting ISIS on Facebook - Breaking the ISIS Brand Counter-Narratives Project*. The International Center for the Study of Violent Extremism (ICSVE). Retrieved from: <https://www.icsve.org/fighting-isis-on-facebook-breaking-the-isis-brand-counter-narratives-project> (01/10/2020).

Speckhard A., Shajkovci A., Bodo L., Fazliu H. (2018). *Bringing Down the Digital Caliphate: A Breaking the ISIS Brand Counter-Narratives Intervention with Albanian*

Speaking Facebook Accounts. The International Center for the Study of Violent Extremism (ICSVE). Retrieved from: <https://www.icsve.org/bringing-down-the-digital-caliphate-a-breaking-the-isis-brand-counter-narratives-intervention-with-albanian-speaking-facebook-accounts> (10/10/2020).

Speckhard A., Shajkovci A., Ahmed M. (2019). Intervening in and Preventing Somali-American Radicalization with Counter Narratives: Testing the Breaking the ISIS Brand Counter Narrative Videos in American Somali Focus Group Settings. *Journal of Strategic Security*, 11, 4: 32. DOI: 10.5038/1944-0472.11.4.1695.

Sundar S.S. (2008). The MAIN Model: A Heuristic Approach to Understanding Technology Effects on Credibility. In Metzger M.J., Flanagin A.J., eds., *Digital Media, Youth, and Credibility*. The John D. & Catherine T. MacArthur Foundation Series on Digital Media and Learning. Cambridge, MA: The MIT Press. DOI: 10.1162/dmal.9780262562324.073. Retrieved from: <https://betterlegalinfo.ca/wp-content/uploads/2019/12/Sundar-paper.pdf> (22/12/2021).

van Eerten J.J., Doosje B. (2020). *Challenging Extremist Views on Social Media*. London: Routledge. DOI: 10.4324/9780429287145.

Wathen C.N., Burkell J. (2002). Believe It or Not: Factors Influencing Credibility on the Web. *Journal of the American Society for Information Science and Technology*, 53, 2: 134. DOI: 10.1002/asi.10016.

Williams M.J. (2020). *Preventing and Countering Violent Extremism: Designing and Evaluating Evidence-Based Programs*. London: Routledge. DOI: 10.4324/9780429441738.