Australian Government
Australian Institute of Criminology

Trends & issues in crime and criminal justice

No. 673

Abstract | This paper reviews available research on how the internet facilitates radicalisation and measures to prevent it. It briefly canvasses evidence on the extent to which the internet contributes to radicalisation broadly, and who is most susceptible to its influence, before delving further into the mechanisms underpinning the relationship between the internet and violent extremism.

High-level approaches to combating internet-facilitated radicalisation, including content removal, account suspensions, reducing anonymity, and counternarrative and education campaigns, are mapped against these mechanisms. This illustrates how these approaches can disrupt radicalisation and assists researchers, policymakers and practitioners to identify potential gaps in existing counterterrorism and countering violent extremism regimes. Research on the implementation and outcomes of these approaches is also summarised.

Understanding and preventing internet-facilitated radicalisation

Heather Wolbers, Christopher Dowling, Timothy Cubitt and Chante Kuhn

The internet has emerged as a central contributor to radicalisation and violent extremist activity (Winter et al. 2020). Facilitating calls for violence in the name of political, religious or other ideologies, the internet serves as a virtual gathering space for like-minded individuals across the globe to discuss, shape and promote violent extremist ideologies, including recruiting people into groups or communities centred on these ideologies. As violent extremism becomes more salient across the internet, the risk that individuals will be radicalised online, recruited into violent extremist groups and inspired or directed by those online to carry out real-world violence also increases.

Consequently, the internet has become a key arena in the struggle against violent extremism. National and international counterterrorism and countering violent extremism policies have increasingly reflected the critical role now played by the internet in radicalisation, emphasising the importance of measures targeting online domains (Australian Government 2022; Montrond et al. 2022; United Nations 2015).

These include the detection and removal of online violent extremist materials, counter-messaging and strategic communications initiatives, media and digital literacy measures to help young people more effectively critique violent extremist messaging, and the fostering of partnerships between governments, media and technology companies.

This paper discusses prevention and intervention approaches aimed at tackling the online drivers of radicalisation. To inform targeted policy, law enforcement and private sector responses, it is critical to understand how the internet contributes to the radicalisation process, how it interacts with other offline risk factors, and who is most susceptible to its influence and why. As such, to guide our discussion of prevention approaches, we first summarise research evidence on the nature and extent of the internet's role in radicalisation, individual vulnerabilities to its influence, and how it contributes to radicalisation and violent extremism.

This paper was informed by literature searches undertaken by the Australian Institute of Criminology's JV Barry Library, which used standard search terms (eg online radicali?ation, preventing violent extremism) to canvass academic and grey literature databases (eg EBSCO, RMIT, ProQuest). We then conducted follow-up searches using citation chaining, and also included additional studies located during further informal searches. As online environments are ever-changing, we limited the review to literature published after 2015, ensuring a focus on relevant and up-to-date information, except where work published prior to 2015 was considered foundational for the field of knowledge. Given the diversity of violent extremist ideologies, we did not focus on any specific ideological background; however, most of the studies reviewed examine Islamist extremism, with some focused on right-wing, issue-specific or lone actor extremism.

Radicalisation

Since the term 'radicalisation' entered public vocabulary in the early 2000s, much academic discourse has discussed its definition (Winter et al. 2020). While various definitions have been offered in research and policy, it is generally understood to be a gradual process, involving multiple drivers, through which individuals accept moral or instrumental justifications for violence in pursuit of a social, economic, political or religious goal that often involves fundamental societal change (della Porta & LaFree 2012; Living Safe Together 2015; Neumann 2013).

A key difference in definitions and conceptualisations of radicalisation, one of particular relevance to this paper, relates to the end point of this process (Macdonald & Whittaker 2019; Neumann 2013). Some definitions emphasise processes of developing violent extremist attitudes and beliefs (ie attitudinal/cognitive radicalisation), while others focus on the lead-up to involvement in acts of violent extremism (ie behavioural radicalisation). While behavioural radicalisation is a primary concern from a public policy standpoint, the development of violent extremist attitudes and beliefs, regardless of whether they manifest in violence, is also relevant (Hardy 2018). The research reviewed in this paper predominantly examines behavioural radicalisation but includes some research on attitudinal radicalisation.

The internet's role in radicalisation

While the majority of radicalised individuals use the internet in some capacity, evidence suggests radicalisation cannot often be solely attributed to online influences (Bastug, Douai & Akca 2020; Gill et al. 2017; Hamid & Ariza 2022; Hollewell & Longpré 2021; Kenyon, Bender & Baker-Beall 2023; Meleagrou-Hitchens, Alexander & Kaderbhai 2017; Mølmen & Ravndal 2021; Whittaker 2022; Winter et al. 2020). There is a growing emphasis on online and offline influences being intertwined, often reinforcing each other (Gaudette, Scrivens & Venkatesh 2022; Gill 2015; Lindekilde, Malthaner & O'Connor 2019; Valentini, Lorusso & Stephan 2020).

Nonetheless, recent studies and systematic reviews have shown that passive and active exposure to violent extremist content online, and engagement with other violent extremists over the internet, are associated with violent extremist attitudes and behaviours (Frissen 2021; Hassan et al. 2018; Pauwels & Schils 2016; Wolfowicz, Hasisi & Weisburd 2022). Moreover, evidence shows significant (if varying) proportions of violent extremist offenders—as many as 60 percent—now either radicalise primarily online or have significant online influences in their radicalisation (Hamid & Ariza 2022; Kenyon, Binder & Baker-Beall 2023).

Susceptibility to internet-facilitated radicalisation

Many people are not susceptible to the influence of content produced by violent extremist groups, and even among those who are drawn to it to some extent, very few go on to engage in acts of violent extremism. At an individual level, those who seek out violent extremist content often have pre-existing risk factors for radicalisation (Mølmen & Ravndal 2021). Furthermore, despite the possibility that exposure to violent extremist content may accelerate radicalisation, in isolation it is unlikely to act as a trigger (Alava, Frau-Meigs & Hassan 2017; Gill et al. 2017; Hassan et al. 2018; Mølmen & Ravndal 2021).

Studies have identified that risk factors for internet-facilitated radicalisation mirror those found in the broader radicalisation literature. These risk factors include:

- various sociodemographic characteristics (being male, young, unemployed or underemployed);
- psychological characteristics (low self-control, personal grievance, certain mental health conditions); and
- contextual characteristics (criminal history or associations with other radicalised persons; Desmarais et al. 2017; LaFree et al. 2018; Wolfowicz et al. 2020).

Social isolation and rejection by peers have emerged as particularly important, likely driving both increased time spent online, which increases exposure opportunities, and an increased vulnerability to radicalisation when violent extremist materials are encountered (Mølmen & Ravndal 2021).

How the internet facilitates radicalisation

A variety of theoretical models have been proposed to explain how radicalisation occurs (eg Borum 2003; Hafez & Mullins 2015; Jensen, Atwell & James 2020; Kruglanski et al. 2018, 2014; McCauley & Moskalenko 2017; Sageman 2004), although comparatively few focus exclusively on online influences or emphasise how the internet contributes to radicalisation (Bastug, Douai & Akca 2020; Mølmen & Ravndal 2021; Neo 2016; Saifudeen 2014; Weimann & Von Knop 2008). These models, along with existing research, point to a number of mechanisms underpinning the relationship between the internet and radicalisation. Critically, these mechanisms are not unique to the online environment. Broadly speaking, the internet accelerates and exacerbates certain elements of the radicalisation process by increasing the accessibility and reach of violent extremists and their messages.

Normalisation and desensitisation

Online environments make vast amounts of violent extremist content quick and easy to access for very large audiences (Neo, Dillon & Khader 2017; Neo et al. 2016). As individuals encounter more of this content, and encounter it more frequently, they can gradually become desensitised to acts of violence and violent extremist ideas, which are normalised and reinforced (Mølmen & Ravndal 2021). Filtering algorithms used by many online platforms that tailor the content presented to people based on their prior activity can further saturate them in online violent extremist content that they may initially have encountered inadvertently or out of curiosity, gradually creating an 'echo chamber' devoid of alternative content (Whittaker 2020). Gradual immersion in online groups and forums populated by those holding violent extremist views can have a similar effect (Atari et al. 2022; Bright 2017; Mølmen & Ravndal 2021).

Persuasiveness

The quality and variety of online violent extremist content can make it far more persuasive than content available offline (Neo et al. 2016). Professionally produced and edited videos, live streamed violent extremist attacks and demonstrations, and even video games have become immersive and emotionally compelling online vehicles for violent extremists to disseminate their ideas, 'perform' violence, and radicalise larger numbers of people (Brzuszkiewicz 2020; Robinson & Whittaker 2021; Scrivens & Conway 2020). Content is also made more persuasive by the adaptability and volume of content, and the currency and regularity of updates, which promote the perception that violent extremists are 'on top of' current events and capable and active in fighting for the cause (Neo et al. 2016).

The accessibility and dynamics of violent extremist groups

Beyond facilitating access to violent extremist content, the internet also brings together those at risk of radicalisation and violent extremists across the world (Conway, Scrivens & McNair 2019; Droogan, Waldek & Blackhall 2018; Hafez & Mullins 2015; Rusumanov 2016; Spalek 2016; Wolfowicz et al. 2020). It allows violent extremists to model or actively convey violent extremist ideas to these at-risk individuals, share violent extremist content with them, and bring larger numbers of them deeper into violent extremist groups and communities.

The ways in which membership of an online violent extremist group or community contribute to the radicalisation process are similar to those in offline groups. Social identity theory posits that belonging to a group impacts perceptions of other groups and one's self-image (Strindberg 2020). Online extremist groups and communities often promote social categorisation by defining clear in-group and out-group identities, which can result in polarisation (ie a sharp division between two groups or ideologies). These communities tend to hold disparaging views of the out-group and exaggerate differences between the in-group and out-group, creating conflict (Mølmen & Ravndal 2021). Perceptions of being marginalised or poorly treated, and ascribing blame to the out-group, can drive endorsement of hostile or violent actions against the out-group (Strindberg 2020). Individuals can be socialised into adopting the in-group identity, and other members or leaders are able to saturate the environment with polarising views and to selectively present narratives and information that promote radicalisation to violent extremism.

Certain processes like socialisation, bonding and peer pressure within online groups and communities of violent extremists can gradually shift inhibitions and create a proclivity for violence (Muro & Wilson 2022). Socialisation can also make individuals feel attached to the group or community, and adopt its norms and values in order to conform and connect with other members. Members can be encouraged to break ties with friends and family who do not conform, further entrenching them in the group and distancing them from prosocial relationships, consequently driving radicalisation (Doosje et al. 2016).

Deindividuation

Once entrenched in an online violent extremist group or community, a group identity may displace the individual identity, which can diffuse moral responsibility and thus contribute to removing normal constraints on behaviour (ie deindividuation; Dalgaard-Nielsen 2008). Deindividuation is further driven by the anonymity offered online, which distances people both from their individual identity and from those they interact with or target, while also bolstering a sense of security (Chang 2008; Neo et al. 2016).

Preventing internet-facilitated radicalisation

A range of approaches to addressing online drivers of radicalisation have been adopted in Australia and internationally. As illustrated in Table 1, these approaches target one or more of the mechanisms discussed above to obstruct the internet's contribution to radicalisation. Mapping these prevention approaches against radicalisation mechanisms in this way can illustrate, at a high level, what a comprehensive national and international regime for countering internet-facilitated radicalisation looks like, and can assist policymakers and practitioners to identify gaps in existing strategies. The following section canvasses each of these approaches and how they work in greater detail, along with research on their effectiveness.

Importantly, evaluations of these approaches are often limited by the use of non-experimental designs, small samples, and challenges inherent in measuring outcomes (Brouillette-Alarie et al. 2022; Carthy et al. 2020; Mastroe & Szmania 2016). Regarding the latter, few studies are able to examine the direct impact of prevention approaches on acts of violent extremism, given how rarely they occur; instead, they examine the impact on intermediate outcomes believed to lead to a reduction in violent extremism (Brouillette-Alarie et al. 2022). Many also draw on the feedback of program providers and stakeholders to gauge effectiveness, which, while informative, introduces significant bias (Brouillette-Alarie et al. 2022). Finally, few studies consider the potential negative or iatrogenic outcomes of these approaches. Any summary of the evidence regarding approaches to combatting radicalisation must acknowledge these limitations.

Table 1: Examples of online methods for preventing radicalisation

Prevention method	How it works	Radicalisation mechanisms targeted
Online content detection and removal	Increases the effort required to disseminate and locate violent extremist content Reduces the risk of individuals inadvertently coming across violent extremist content	Normalisation and reinforcement of violent extremism Desensitisation to violent extremism
Suspending extremist accounts	Disrupts formation of violent extremist communities Limits potential for connections with violent extremists Inhibits the creation and dissemination of violent extremist content	Accessibility of violent extremist groups Normalisation and reinforcement of violent extremism Desensitisation to violent extremism Peer pressure and socialisation within violent extremist groups Deindividuation Polarisation
Reducing anonymity	Increases the risk involved in engaging with extremist content and communities	Deindividuation
Counternarratives, alternative narratives and strategic communications	Directly critique the narratives offered in extremist ideologies Offer competing narratives on events or phenomena to those offered in extremist ideologies	Normalisation and reinforcement of violent extremism Desensitisation to violent extremism Persuasiveness of extremist messaging Polarisation
Education in civics and critical media consumption	Builds resilience to radical ideas Promotes doubt and questioning of radical ideologies	Normalisation of violent extremism Persuasiveness of extremist messaging

Content detection and removal

Detection and removal of violent extremist content makes this content less accessible, increasing the effort required by at-risk individuals to find and engage with it, and reducing the risk of inadvertently viewing it (Hardy 2022). Reducing the accessibility of content also obstructs the processes of normalisation and reinforcement of violent extremist ideas.

Implementing content detection and removal schemes has several practical challenges. Due to the nature of the internet, violent extremist content can rapidly proliferate beyond the capacity of reasonable control measures. For example, following the Christchurch attack, which was live streamed by the perpetrator, platforms like Facebook and YouTube were unable to stop the rapid spread of the video as millions of reproductions were uploaded while users circumvented systems in place to prevent this (Dwoskin & Timberg 2019).

Moderation systems tend to include both automated and human-driven methods to identify and remove online violent extremist content (Bradford et al. 2019; Díaz & Hecht-Felella 2021). While automated methods outperform human-driven approaches, aspects of human review will continue to be needed due to biases and errors, making online content detection and removal a time- and labour-intensive process (eg Hall et al. 2020). Because of these challenges, many platforms lack a robust moderation regime, creating vulnerabilities that can allow the spread of violent extremist content on these platforms, and potentially onto others. Nonetheless, content removal schemes have demonstrated some success at removing vast quantities of violent extremist content. For example, Facebook removed 9.4 million pieces of Islamist extremism related content between April and June 2018 (Counter Terrorism Policing 2018). The amount of violent extremist content that went undetected, however, remains unknown, meaning it is difficult to ascertain the effectiveness of moderation.

Suspending accounts

The suspension of online accounts that produce or disseminate violent extremist messaging is also used to stem the flow of content and disrupt the formation and growth of violent extremist groups and communities (Hafez & Mullins 2015; Neo, Dillon & Khader 2017). Suspending accounts used by violent extremists can have a number of benefits. For example, some studies have found that suspending accounts supportive of Islamic State greatly affected the terrorist organisation's ability to develop and maintain communities online (Berger & Perez 2016; Conway, Scrivens & McNair 2019). Further, Chandrasekharan and colleagues (2017) found that active users on hate-based Reddit subforums which were shut down became active on other parts of Reddit, but their expressions of hate, misogyny and racism decreased. However, different platforms treat violent extremist content with different levels of seriousness, and banned users may move to other platforms where their potential audience is smaller but they are able to operate with fewer restrictions and a lower risk of having their account suspended (Ali et al. 2021). Overall, while specific platforms can suspend accounts, without broader coordination and consistency, problematic individuals can move from site to site, establishing a presence on platforms where moderation standards are lower.

Unsurprisingly, the effectiveness of measures that reduce the availability of violent extremist content, either through content removal or account suspension, appear contingent on successful collaboration among stakeholders across civil society and the private sector rather than government alone (Aly, Balbi, & Jacques 2015; Briggs & Feve 2013; Brown & Marway 2018; Dalgaard-Nielsen 2016; Gielen 2019). With this in mind, partnerships between government, technology and online media companies have become more common as they attempt to harmonise, coordinate and maximise the reach of these measures. For example, the Global Internet Forum to Counter Terrorism (2022) maintains and shares a database of hashes (digital fingerprints) for extremist content that governments and industry can use to identify and block problematic content online.

Reducing anonymity

Another measure for decreasing the risk of radicalisation to violent extremism online involves reducing the anonymity of users. This can include enforcing identity checks or implementing multifactor authentication for account creation (Hardy 2022). As discussed, user anonymity bolsters a sense of security on online platforms where individuals feel the risk of identification, detection and apprehension is low (Neo et al. 2016). It is also associated with deindividuation, in which people lose their sense of individuality and instead align their identity with a group, limiting the sense of responsibility for antisocial behaviours (Chang 2008). Anonymity online can weaken inhibitions and encourage people to behave and engage with content in ways they would not in the real world (Demetriou & Silke 2003).

Concerns have been raised, however, that reducing online anonymity would negatively affect online privacy and freedom of speech and lead to discrimination (Luca 2022; Scott 2004). Critically, the internet is a key outlet for disseminating information that has been suppressed by authoritarian regimes, and empowering resistance against corrupt or criminal practices by individuals, companies (eg the Panama Papers) or governments (eg Arab Spring). As such, it is important to be aware of how measures to reduce online anonymity and track individual identities might be misused to suppress the free expression of ideas and information.

Counternarratives, alternative narratives and strategic communications

Implementing counternarratives, alternative narratives and strategic communications involves government or private industry disseminating messages in order to foster relationships between government and communities, discredit the ideologies and actions of violent extremists, offer alternative points of view, and directly contradict violent extremist messaging. Counternarrative campaigns, for example, discredit violent extremist groups and their ideologies by deconstructing and demystifying their messages to demonstrate lies, hypocrisy and inconsistencies (Speckhard, Shajkovci & Ahmed 2018). These approaches counteract the effects of engaging with violent extremist content, such as scepticism of mainstream views and the subsequent development of problematic ideologies, and neutralise provocations by making violent extremist messages less appealing. Similarly, echo chamber effects and the adoption of polarising views can be decreased by making alternative ideas more widely and readily available, impeding the entrenchment and validation of biased perspectives. Examples of Australian campaigns include the All Together Now application, which aims to build awareness of racism and debunk far-right extremist ideology and associated misinformation campaigns (All Together Now 2022). Its developer reports that users have an increased awareness of racism and are more likely to speak up against racism.

While counternarrative campaigns are widely discussed and advocated for, there is often uncertainty around whether they are reaching or resonating with their intended audience (Alava, Frau-Meigs & Hassan 2017; Bélanger et al. 2020; Brouillette-Alarie et al. 2022; Carthy et al. 2020; Rosand & Winterbotham 2019). Many evaluations of campaigns analyse reach, views and engagement, which does not offer insight into actual impacts on the audience's attitudes or behaviour (Helmus & Klein 2018). Nevertheless, evidence suggests that these campaigns may be more effective with people in the early stages of radicalisation than with those who already hold ingrained violent extremist views (Carthy et al. 2020). Further, the persuasiveness of these campaigns is maximised when they are delivered by those who are seen as 'one of them' (Braddock & Horgan 2016; Freear & Glazzard 2020). Conversely, counternarrative campaigns, if undertaken too aggressively and with individuals who already hold entrenched violent extremist views, can backfire and drive radicalisation instead of reversing or preventing it (eg Bélanger et al. 2020; Bodine-Baron et al. 2020).

Education

Finally, some approaches aim to reduce the risk of radicalisation occurring online through educational measures. Education may build awareness of democracy, pluralism and peaceful ideas to promote mainstream views and mitigate the rejection of these ideals and the normalisation of violent extremist beliefs (Alava, Frau-Meigs & Hassan 2017; Neumann 2013). Prior Australian programs have primarily aimed to counter the jihadi interpretation of Islam among youth and convicted terrorists (Akbarzadeh 2013). Such approaches have shown some success but have been criticised for not addressing the underlying causes of radicalisation, for not addressing other forms of violent extremism, and for equating to state interference in religion through sponsorship of an 'acceptable' version of Islam.

Education approaches may also focus on bolstering awareness of online risks, and developing digital and media literacy skills in the general community. Such skills promote critical consumption of digital media to build resilience to violent extremist messages (Alava, Frau-Meigs & Hassan 2017; Davies 2018; Schmitt et al. 2018). In practice, educational approaches develop an individual's ability to critically consume violent extremist content, and to interpret and make informed decisions about how to engage with its messages, ultimately reducing susceptibility to the influence of this content (Neo, Dillon & Khader 2017).

Conclusion

Exposure to violent extremist content online and engagement with violent extremists over the internet is associated with the development of violent extremist attitudes and behaviours. Research also highlights the different mechanisms through which the internet drives radicalisation. Of course, the extent to which these mechanisms operate on someone depends on co-occurring individual vulnerabilities and real-world influences. Nevertheless, the internet's role in radicalisation to violent extremism is currently, and will continue to be, an important area of focus for those tasked with addressing this pervasive threat, and requires a comprehensive and multifaceted approach.

Importantly, addressing online drivers of radicalisation is just one part of a comprehensive countering violent extremism and counterterrorism strategy. Given the overlap and interaction between online and real-world influences, online measures must work in tandem with broader approaches. This is consistent with current countering violent extremism strategies, which tend to include initiatives that span primary, secondary and tertiary prevention, to target the societal, community, interpersonal and individual elements of radicalisation, and to address multiple forms of violent extremism (Ambrozik 2019; Kundnani & Hayes 2018). Measures aimed at addressing online drivers of radicalisation are best situated within this broad regime to maximise their effectiveness.

References

URLs correct as at May 2023

Akbarzadeh S 2013. Investing in mentoring and educational initiatives: The limits of de-radicalisation programmes in Australia. *Journal of Muslim Minority Affairs* 33(4): 451–463. https://doi.org/10.1080/13602004.2013.866347

Alava S, Frau-Meigs D & Hassan G 2017. *Youth and violent extremism on social media: Mapping the research.* UNESCO Publishing. https://unesdoc.unesco.org/ark:/48223/pf0000260382

Ali S et al. 2021. *Understanding the effect of deplatforming on social networks.* Proceedings of the 13th ACM Web Science Conference 2021. Online event, UK: Association for Computing Machinery: 187–195. https://doi.org/10.1145/3447535.3462637

All Together Now 2022. Everyday racism. https://alltogethernow.org.au/our-work/everyday-racism/

Aly A, Balbi A & Jacques C 2015. Rethinking countering violent extremism: Implementing the role of civil society. *Journal of Policing, Intelligence and Counter Terrorism* 10(1): 3–13. https://doi.org/10.1080/18335330.2015.1028772

Ambrozik C 2019. Countering violent extremism globally: A new global CVE dataset. *Perspectives on Terrorism* 13(5): 102–111. https://www.jstor.org/stable/26798581

Atari M et al. 2022. Morally homogeneous networks and radicalism. *Social Psychological and Personality Science* 13(6): 999–1009. https://doi.org/10.1177/19485506211059329

Australian Government 2022. *Safeguarding our community together: Australia's counter-terrorism strategy.* Canberra: Department of Home Affairs. https://www.nationalsecurity.gov.au/what-australia-is-doing/a-national-approach/australias-counter-terrorism-strategies

Bastug MF, Douai A & Akca D 2020. Exploring the "demand side" of online radicalization: Evidence from the Canadian context. *Studies in Conflict & Terrorism* 43(7): 616–637. https://doi.org/10.1080/1057610X.2018.1494409

Bélanger JJ, Nisa CF, Schumpe BM, Gurmu T, Williams MJ & Putra IE 2020. Do counter-narratives reduce support for ISIS? Yes, but not for their target audience. *Frontiers in Psychology* 11: 1–11. https://doi.org/10.3389/fpsyg.2020.01059

Berger JM & Perez H 2016. *The Islamic State's diminishing returns on Twitter: How suspensions are limiting the social networks of English-speaking ISIS supporters*. George Washington University. https://extremism.gwu.edu/isis-online-reports

Bodine-Baron E, Marrone JV, Helmus TC & Schlang D 2020. *Countering violent extremism in Indonesia: Using an online panel survey to assess a social media counter-messaging campaign.* Santa Monica, California: RAND Corporation. https://www.rand.org/pubs/research_reports/RRA233-1.html

Borum R 2003. Understanding the terrorist mind-set. *Crime & Justice International* 19(77): 28–30. https://www.ojp.gov/ncjrs/virtual-library/abstracts/understanding-terrorist-mindset

Braddock K & Horgan J 2016. Towards a guide for constructing and disseminating counternarratives to reduce support for terrorism. *Studies in Conflict & Terrorism* 39(5): 381–404. https://doi.org/10.1080/1057610X.2015.1116277

Bradford B et al. 2019. *Report of the Facebook Data Transparency Advisory Group*. New Haven: Justice Collaboratory: Yale Law School. https://www.justicehappenshere.yale.edu/reports/report-of-the-facebook-data-transparency-advisory-group

Briggs R & Feve S 2013. *Review of programs to counter narratives of violent extremism: What works and what are the implications for government?* London: Institute for Strategic Dialogue. https://www.publicsafety.gc.ca/lbrr/archives/cn28580-eng.pdf

Bright J 2017. *Explaining the emergence of echo chambers on social media: The role of ideology and extremism*. https://ssrn.com/abstract=2839728

Brouillette-Alarie S et al. 2022. Systematic review on the outcomes of primary and secondary prevention programs in the field of violent radicalization. *Journal for Deradicalization* 30: 117–168. https://journals.sfu.ca/jd/index.php/jd/article/view/577

Brown K & Marway H 2018. *Preventing radicalisation to terrorism and violent extremism: Delivering counter- or alternative narratives*. Radicalisation Awareness Network

Brzuszkiewicz S 2020. Jihadism and far-right extremism: Shared attributes with regard to violence spectacularisation. *European View* 19(1): 71–79. https://doi.org/10.1177/1781685820915972

Carthy SL, Doody CB, Cox K, O'Hora D & Sarma KM 2020. Counter-narratives for the prevention of violent radicalisation: A systematic review of targeted interventions. *Campbell Systematic Reviews* 16(3): e1106. https://doi.org/10.1002/cl2.1106

Chandrasekharan E, Pavalanathan U, Srinivasan A, Glynn A, Eisenstein J & Gilbert E 2017. You can't stay here: The efficacy of Reddit's 2015 ban examined through hate speech. *Proceedings of the ACM Human-Computer Interaction* 1(31): 1–22. https://doi.org/10.1145/3134666

Chang J 2008. The role of anonymity in deindividuated behavior: A comparison of deindividuation theory and the social identity model of deindividuation effect. *The Pulse* 6(1): 1–8

Conway M, Scrivens R & McNair L 2019. *Right-wing extremists' persistent online presence: History and contemporary trends.* International Centre for Counter-Terrorism Policy Brief. The Hague: International Centre for Counter-Terrorism. https://www.icct.nl/publication/right-wing-extremists-persistent-online-presence-history-and-contemporary-trends

Counter Terrorism Policing 2018. Together we're tackling online terrorism. *Counter Terrorism Policing News*, December 2018. https://www.counterterrorism.police.uk/together-were-tackling-online-terrorism/

Dalgaard-Nielsen A 2016. Countering violent extremism with governance networks. *Perspectives on Terrorism* 10(6): 135–139. https://www.jstor.org/stable/26297713

Dalgaard-Nielsen A 2008. *Studying violent radicalization in Europe II: The potential contribution of socio-psychological and psychological approaches*. Working Paper no 2008/3. Copenhagen: Danish Institute for International Studies. https://www.econstor.eu/bitstream/10419/84593/1/DIIS2008-03.pdf

Davies L 2018. *Review of educational initiatives in counter-extremism internationally: What works?* Gothenburg: Segerstedt Institute University of Gothenburg. https://gupea.ub.gu.se/handle/2077/66726

della Porta D & LaFree G 2012. Processes of radicalization and deradicalization. *International Journal of Conflict and Violence* 6(1): 4–10

Demetriou C & Silke A 2003. A criminological internet 'sting': Experimental evidence of illegal and deviant visits to a website trap. *British Journal of Criminology* 43(1): 213–222. https://doi.org/10.1093/bjc/43.1.213

Desmarais SL, Simons-Rudolph J, Brugh CS, Schilling E & Hoggan C 2017. The state of scientific knowledge regarding factors associated with terrorism. *Journal of Threat Assessment and Management* 4(4): 180–209. https://doi.org/10.1037/tam0000090

Díaz Á & Hecht-Felella L 2021. *Double standards in social media content moderation.* New York: Brennan Center for Justice at New York University School of Law. https://www.brennancenter.org/our-work/research-reports/double-standards-social-media-content-moderation

Doosje B, Moghaddam FM, Kruglanski AW, De Wolf A, Mann L & Feddes AR 2016. Terrorism, radicalization and de-radicalization. *Current Opinion in Psychology* 11: 79–84. https://doi.org/10.1016/j.copsyc.2016.06.008

Droogan J, Waldek L & Blackhall R 2018. Innovation and terror: An analysis of the use of social media by terror-related groups in the Asia Pacific. *Journal of Policing, Intelligence and Counter Terrorism* 13(2): 170–184. https://doi.org/10.1080/18335330.2018.1476773

Dwoskin E & Timberg C 2019. Christchurch mosque shootings: Inside YouTube's struggles to shut down video – and the humans who outsmarted its systems. *Washington Post*, 18 March. https://www.washingtonpost.com/technology/2019/03/18/inside-youtubes-struggles-shut-down-video-new-zealand-shooting-humans-who-outsmarted-its-systems/

Freear M & Glazzard A 2020. Preventive communication: Emerging lessons from participative approaches to countering violent extremism in Kenya. *The RUSI Journal* 165(1): 90–106. https://doi.org/10.1080/03071847.2020.1734316

Frissen T 2021. Internet, the great radicalizer? Exploring relationships between seeking for online extremist materials and cognitive radicalization in young adults. *Computers in Human Behavior* 114: 106549. https://doi.org/10.1016/j.chb.2020.106549

Gaudette T, Scrivens R & Venkatesh V 2022. The role of the internet in facilitating violent extremism: Insights from former right-wing extremists. *Terrorism and Political Violence* 34(7): 1339–1356. https://doi.org/10.1080/09546553.2020.1784147

Gielen A 2019. Countering violent extremism: A realist review for assessing what works, for whom, in what circumstances, and how? *Terrorism and Political Violence* 31(6): 1149–1167. https://doi.org/10.1080/09546553.2017.1313736

Gill P 2015. *Lone actor terrorists: A behavioural analysis*. Oxford: Routledge

Gill P, Corner E, Conway M, Thornton A, Bloom M & Horgan J 2017. Terrorist use of the internet by the numbers: Quantifying behaviors, patterns, and processes. *Criminology & Public Policy* 16(1): 99–117. https://doi.org/10.1111/1745-9133.12249

Global Internet Forum to Counter Terrorism (GIFCT) 2022. GIFCT's Hash-Sharing Database. https://gifct.org/hsdb/

Hafez M & Mullins C 2015. The radicalization puzzle: A theoretical synthesis of empirical approaches to homegrown extremism. *Studies in Conflict & Terrorism* 38(11): 958–975. https://doi.org/10.1080/1057610X.2015.1051375

Hall M, Logan M, Ligon GS & Derrick DC 2020. Do machines replicate humans? Toward a unified understanding of radicalizing content on the open social web. *Policy & Internet* 12(1): 109–138. https://doi.org/10.1002/poi3.223

Hamid N & Ariza C 2022. *Offline versus online radicalization: Which is the bigger threat? Tracing outcomes of 439 jihadist terrorists between 2014–2021 in 8 Western countries*. Global Network on Extremism and Technology. https://gnet-research.org/2022/02/21/offline-versus-online-radicalisation-which-is-the-bigger-threat/

Hardy K 2022. A crime prevention framework for CVE. *Terrorism and Political Violence* 34(3): 633–659. https://doi.org/10.1080/09546553.2020.1727450

Hardy K 2018. Comparing theories of radicalisation with countering violent extremism policy. *Journal for Deradicalization* 15: 76–110. https://journals.sfu.ca/jd/index.php/jd/article/view/150

Hassan G et al. 2018. Exposure to extremist online content could lead to violent radicalization: A systematic review of empirical evidence. *International Journal of Developmental Science* 12(1–2): 71–88. https://doi.org/10.3233/DEV-170233

Helmus TC & Klein K 2018. *Assessing outcomes of online campaigns countering violent extremism: A case study of the Redirect Method.* Santa Monica, CA: RAND Corporation. https://www.rand.org/pubs/research_reports/RR2813.html

Hollewell GF & Longpré N 2021. Radicalization in the social media era: Understanding the relationship between self-radicalization and the internet. *International Journal of Offender Therapy and Comparative Criminology* 66(8): 896–913. https://doi.org/10.1177/0306624X211028771

Jensen MA, Atwell AA & James PA 2020. Radicalization to violence: A pathway approach to studying extremism. *Terrorism and Political Violence* 32(5): 1067–1090. https://doi.org/10.1080/09546553.2018.1442330

Kenyon J, Binder JF & Baker-Beall C 2023. Online radicalization: Profile and risk analysis of individuals convicted of extremist offences. *Legal and Criminological Psychology* 28(1): 74–90. https://doi.org/10.1111/lcrp.12218

Kruglanski AW, Gelfand MJ, Bélanger JJ, Shevland A, Hetiarachcho M & Gunaratna R 2014. The psychology of radicalization and deradicalization: How significance quest impacts violent extremism. *Political Psychology* 35(1): 69–93. https://doi.org/10.1111/pops.12163

Kruglanski AW, Jasko K, Webber D, Chernikova M & Molinario E 2018. The making of violent extremists. *Review of General Psychology* 22(1): 107–120. https://doi.org/10.1037/gpr0000144

Kundnani A & Hayes B 2018. The globalisation of countering violent extremism policies: Undermining human rights, instrumentalising civil society. Amsterdam: Transnational Institute. https://www.tni.org/en/publication/the-globalisation-of-countering-violent-extremism-policies

LaFree G, Jensen MA, James PA & Safer-Lichtenstein A 2018. Correlates of violent political extremism in the United States. *Criminology* 52(2): 233–268. https://doi.org/10.1111/1745-9125.12169

Lindekilde L, Malthaner S & O'Connor F 2019. Peripheral and embedded: Relational patterns of lone-actor terrorist radicalization. *Dynamics of Asymmetric Conflict: Pathways toward Terrorism and Genocide* 12(1): 20–41. https://doi.org/10.1080/17467586.2018.1551557

Living Safe Together 2015. *Preventing violent extremism and radicalisation in Australia*. Canberra: Australian Government Attorney-General's Department. https://www.livingsafetogether.gov.au/resources

Luca S 2022. In defense of online anonymity. *Wall Street Journal*, 17 June. https://www.wsj.com/articles/the-value-of-online-anonymity-11655473116

Macdonald S & Whittaker J 2019. Online radicalization: Contested terms and conceptual clarity. In JR Vacca (ed), *Online terrorist propaganda, recruitment, and radicalization*. Boca Raton: CRC Press. https://cronfa.swan.ac.uk/Record/cronfa45970

Mastroe C & Szmania S 2016. *Surveying CVE metrics in prevention, disengagement and deradicalization programs*. Report to the Office of University Programs, Science and Technology Directorate. College Park, MD: Department of Homeland Security. https://www.start.umd.edu/publication/surveying-cve-metrics-prevention-disengagement-and-de-radicalization-programs

McCauley C & Moskalenko S 2017. Understanding political radicalization: The two-pyramids model. *American Psychologist* 72(3): 205–216. https://doi.org/10.1037/amp0000062

Meleagrou-Hitchens A, Alexander A & Kaderbhai N 2017. The impact of digital communications technology on radicalization and recruitment. *International Affairs* 93(5): 1233–1249. https://doi.org/10.1093/ia/iix103

Mølmen GN & Ravndal JA 2021. Mechanisms of online radicalisation: how the internet affects the radicalisation of extreme-right lone actor terrorists. *Behavioral Sciences of Terrorism and Political Aggression.* https://doi.org/10.1080/19434472.2021.1993302

Montrond A, Ekström A, Nielson R, Hadji-Janev M & Savoia E 2022. Comparative analysis of CT/CVE policies: USA, Canada, United Kingdom, Sweden, and North Macedonia. *Homeland Security Affairs* 18(1). https://doi.org/10.3390/proceedings2021077006

Muro D & Wilson T (eds) 2022. *Contemporary terrorism studies*. Oxford University Press. https://doi.org/10.1093/hepl/9780198829560.001.0001

Neo LS 2016. An internet-mediated pathway for online radicalisation: RECRO. In M Khader, LS Neo, G Ong, ET Mingyi & J Chin (eds), *Combating violent extremism and radicalization in the digital era*. Hershey, PA: IGI Global: 197–224. https://doi.org/10.4018/978-1-5225-0156-5.ch011

Neo LS, Dillon L & Khader M 2017. Identifying individuals at risk of being radicalised via the internet. *Security Journal* 30: 1112–1133. https://doi.org/10.1057/s41284-016-0080-z

Neo LS, Dillon L, Shi P, Tan J, Wang Y & Gomes D 2016. Understanding the psychology of persuasive violent extremist online platforms. In M Khader, LS Neo, G Ong, ET Mingyi & J Chin (eds), *Combating violent extremism and radicalization in the digital era*. Hershey, PA: IGI Global: 1–15. https://doi.org/10.4018/978-1-5225-0156-5.ch001

Neumann P 2013. Options and strategies for countering online radicalization in the United States. *Studies in Conflict & Terrorism* 36(6): 431–459. https://doi.org/10.1080/1057610X.2013.784568

Pauwels L & Schils N 2016. Differential online exposure to extremist content and political violence: Testing the relative strength of social learning and competing perspectives. *Terrorism and Political Violence* 28(1): 1–29. https://doi.org/10.1080/09546553.2013.876414

Robinson N & Whittaker J 2021. Playing for hate? Extremism, terrorism, and videogames. *Studies in Conflict and Terrorism*. Advance online publication. https://doi.org/10.1080/1057610X.2020.1866740

Rosand E & Winterbotham E 2019. Do counter-narratives actually reduce violent extremism? The Brookings Institution. https://www.brookings.edu/blog/order-from-chaos/2019/03/20/do-counter-narratives-actually-reduce-violent-extremism/

Rusumanov V 2016. The use of the internet by terrorist organizations. *Information & Security* 34(2): 137–150. https://doi.org/10.11610/isij.3409

Sageman M 2004. *Understanding terror networks*. Philadelphia: University of Pennsylvania Press. https://doi.org/10.9783/9780812206791

Saifudeen OA 2014. The cyber extremism orbital pathways model. Singapore: S Rajaratnam School of International Studies, Nanyang Technological University

Schmitt JB, Rieger D, Ernst J & Roth H 2018. Critical media literacy and Islamist online propaganda: The feasibility, applicability and impact of three learning arrangements. *International Journal of Conflict and Violence* 12: 1–19. https://doi.org/10.4119/UNIBI/ijcv.642

Scott CR 2004. Benefits and drawbacks of anonymous online communication: Legal challenges and communicative recommendations. *Free Speech Yearbook* 41(1): 127–141. https://doi.org/10.1080/08997225.2004.10556309

Scrivens R & Conway M 2020. The roles of 'old' and 'new' media tools and technologies in the facilitation of violent extremism and terrorism. In R Leukfeldt & TJ Holt (eds), *The human factor of cybercrime*. Abingdon: Routledge. https://doi.org/10.4324/9780429460593-13

Spalek B 2016. Radicalisation, de-radicalisation and counter-radicalisation in relation to families: Key challenges for research, policy and practice. *Security Journal* 29(1): 39–52. https://doi.org/10.1057/sj.2015.43

Speckhard A, Shajkovci A & Ahmed M 2018. Intervening in and preventing Somali-American radicalization with counter narratives. *Journal of Strategic Security* 11(4): 32–71. https://doi.org/10.5038/1944-0472.11.4.1695

Strindberg A 2020. *Social identity theory and the study of terrorism and violent extremism*. Sweden: Swedish Defence Research Agency

United Nations 2015. *Plan of action to prevent violent extremism*. New York: United Nations. https://www.un.org/counterterrorism/plan-of-action-to-prevent-violent-extremism

Valentini D, Lorusso AM & Stephan A 2020. Onlife extremism: Dynamic integration of digital and physical spaces in radicalization. *Frontiers in Psychology* 11: 524. https://doi.org/10.3389/fpsyg.2020.00524

Weimann G & Von Knop K 2008. Applying the notion of noise to countering online terrorism. *Studies in Conflict & Terrorism* 31(10): 883–902. https://doi.org/10.1080/10576100802342601

Whittaker J 2022. Rethinking online radicalization. *Perspectives on Terrorism* 16(4): 27–40.

Whittaker J 2020. Online echo chambers and violent extremism. In SM Khasru (ed), *The digital age, cyber space, and social media: The challenges of security & radicalization*. Dhaka: IPAG: 129–150

Winter C, Neumann P, Meleagrou-Hitchens A, Ranstorp M, Vidino L & Fürst J 2020. Online extremism: Research trends in internet activism, radicalization, and counter-strategies. *International Journal of Conflict and Violence* 14(2): 1–20. https://doi.org/10.4119/ijcv-3809

Wolfowicz M, Hasisi B & Weisburd D 2022. What are the effects of different elements of media on radicalization outcomes? A systematic review. *Campbell Systematic Reviews* 18(2): e1244. https://doi.org/10.1002/cl2.1244

Wolfowicz M, Litmanovitz Y, Weisburd D & Hasisi B 2020. A field-wide systematic review and meta-analysis of putative risk and protective factors for radicalization outcomes. *Journal of Quantitative Criminology* 36(3): 407–447. https://doi.org/10.1007/s10940-019-09439-4

Dr Heather Wolbers is a Senior Research Analyst at the Australian Institute of Criminology.

Dr Christopher Dowling is a Research Manager at the Australian Institute of Criminology.

Dr Timothy Cubitt is a Principal Research Analyst at the Australian Institute of Criminology's Serious and Organised Crime Research Laboratory.

Chante Kuhn is a former Intern at the Australian Institute of Criminology.

General editor, *Trends & issues in crime and criminal justice* series: Dr Rick Brown, Deputy Director, Australian Institute of Criminology. Note: *Trends & issues in crime and criminal justice* papers are peer reviewed. For a complete list and the full text of the papers in the *Trends & issues in crime and criminal justice* series, visit the AIC website: www.aic.gov.au

ISSN 1836-2206 (Online) ISBN 978 1 922877 02 4 (Online)

https://doi.org/10.52922/ti77024

GPO Box 1936
Canberra ACT 2601, Australia

Tel: 02 6268 7166

Disclaimer: This research paper does not necessarily reflect the policy position of the Australian Government

www.aic.gov.au