UK Parliament POST

POSTNOTE

Number 622 May 2020

Online Extremism

Extremism is possible in any ideology, including (but not limited to) politics and religion. Extremism can affect mental well-being, amplify hostility, and threaten democratic debate. This POSTnote focuses on how extremism manifests online, consequences of exposure, and potential countermeasures.

Overview

- Extremist content may be found on mainstream sites and 'alt-tech' platforms that have been created or co-opted for the unconventional needs of specific users.
- The Internet may facilitate extremism in multiple ways, including recruitment, socialisation and mobilisation.
- Countering online extremism requires a coordinated approach. Methods include content removal and social interventions.
- A key challenge is identifying responsibilities for online content regulation. The 2019 Online Harms White Paper proposed the appointment of an independent regulator.

Background

Extremism lacks a clear definition, which has contributed to difficulty in regulating it.[1,2] The UK Government characterises extremism as "opposition to fundamental values, including democracy, the rule of law, individual liberty, and respect and tolerance for different faiths and beliefs".[3] However, this definition lacks the legal precision for extremism to be grounded in UK law.[4] When the Commission for Countering Extremism (the CCE, a Home Office advisory agency) surveyed over 2,500 members of the public, 75% thought this definition was unhelpful.[5] Among the concerns expressed were that values are not clearly defined, and that subjective interpretation could undermine democracy because people may be labelled extreme for having preferences outside of the mainstream.[6]

The CCE suggests that hateful extremism includes any behaviours that: incite or amplify hate; make a 'moral case' for violence; spread hostile or supremacist beliefs against a particular group; or cause harm to individuals, communities or wider society.[5] Some stakeholders dispute the term 'hateful extremism', arguing that hate speech and extremism are separate societal issues.[7] However, most agree that a narrower definition of extremism is needed for policies to be effective.[8] Extremism is not the same as terrorism. Terrorism is violence (or threat of it) used to intimidate the public or advance a particular cause.[9] While terrorism is a tactic, extremism is a belief system.[10] However, exposure to extremism can encourage an individual's support for terrorist tactics (a process known as radicalisation).[11] The CCE report that 56% of the public agree that a lot more should be done to counter extremism online.[5] In 2019, the UK Government put forward proposals to address online extremism in the Online Harms White Paper.[12]

Extremism can often be attributed to a combination of online and offline interactions, where the Internet acts as an enabler rather than the root cause.[13–17] Social media content can be posted instantly by anyone without verification or external editorial control, so information can be produced rapidly and disseminated widely (POSTnote 559).[18] Contributors can consciously evade detection by using multiple accounts, and there are inconsistencies in how content is moderated across platforms.[12] UK Government, public sector practitioners and industry stakeholders have called for a coordinated response to address the range of social and technological challenges around extremism.[19]

This POSTnote outlines how the online environment can be used for extremist purposes, how exposure to online extremism can influence people, and potential strategies to counter online extremist content.

The online environment

The Internet allows people to reach global audiences quickly, cheaply and easily.[20] This creates positive opportunities for individuals and organisations. However, the global reach of the Internet poses challenges for national governments attempting to safeguard citizens from harm online (Box 1). Technology is continuously evolving, and extremists tend to be early adopters of new tools in order to exploit this.[21] As a result, it is challenging for authorities and regulators to keep up.[22]

The Parliamentary Office of Science and Technology, Westminster, London SW1A 0AA
02072192840 post@parliament.uk parliament.uk/post @POST_UK

Online platforms include mainstream social media sites as well as 'alt-tech' platforms, which replicate the functions of mainstream social media but have been created or co-opted for the unconventional needs of specific users.[23] Both types of platforms present challenges for addressing online extremism.

Mainstream platforms include Facebook, Twitter and YouTube. The advantage of uploading extremist content to mainstream platforms is that content can reach a large audience. During the first half of 2019, Twitter removed over 100,000 accounts promoting terrorism, and locked or suspended over 50,000 accounts for breaching violent extremism policy guidelines.[24] The scale of online extremism is unknown, as estimates may be amplified by fake accounts, which include multiple accounts created by the same person and 'bots' (computer programmes that impersonate humans and post automated content).[25]

Researchers have suggested that YouTube recommendation algorithms, which can filter or prioritise content based on user engagement, can cause a radicalisation pipeline where users migrate to more extreme content that they might not otherwise have been exposed to.[26] The aim of most online platforms is to keep users engaged, and algorithms learn users' interests and biases by exposing them to content and monitoring their response. This can inadvertently create incentives to show users extreme content, as people are more likely to click and comment on extreme headlines.[27] However, other researchers have argued that it is impossible to separate the influence of algorithms from conscious human decisions.[28]

'Alt-tech' platforms hosting extremist content may include:

- Platforms that have been 'hijacked' for extremist purposes but actively cooperate with the authorities.[29,30] Messaging app Telegram is used by extremists because content is encrypted (POSTbrief 19). Telegram worked with the European Union Internet Referral Unit to improve tools for users to flag terrorist content.[31]
- Platforms that tolerate extremist content under promotion of free speech. [29,30] Online forums such as Gab allow users to post potentially extremist content with little repercussion or official moderation.[29,30]
- Self-hosted platforms that were built specifically for extremist content and are owned by extremist groups.[32,33]

Exposure to extremism

Assessing the scale of exposure to extremist content is difficult. People may access it covertly, unknowingly, or be unsure whether they should report it (especially as users differ in what they consider extreme). The CCE survey found that 25% of the UK general public had witnessed extremism online.[5] A study of 706 members of the public found that around 7% had specifically searched for extremist content online.[34] The context of exposure is also likely to matter, with research suggesting that individuals are more likely to engage with extremist content if they are directed to it by friends or family.[35] The link between viewing extremist content and radicalisation or inciting violence is not well understood.[36] Research often looks at prior exposure to online extremism of those who have been radicalised, but there is less research on those who do not develop extreme views following exposure.[16]

Box 1: International cooperation

The majority of content accessed by UK citizens is not hosted in the UK (for example, most mainstream social media platforms are based in the US).[37,38] Countries have different laws and attitudes to regulation, censorship and protection of free speech online.[39] Cooperation between nations is required to enact national legislation.[40] However, a lack of shared international definitions and concepts (such as 'extremism' or 'free speech') makes it difficult to coordinate responses. In 2018, Germany introduced legislation that places sanctions on social media companies that fail to remove illegal terrorist content and hate speech within 24 hours.[41–43] The approach had an approval rate of 87% with German voters.[42] However, some stakeholders have expressed concerns that the approach could set a precedent for some nations to use it for state censorship.[44] Mainstream platforms may have terms and conditions that are more stringent than UK legislation, in order to operate globally and comply with the most stringent international legislation.[45]

The Internet gives people anonymity (which can make users feel disinhibited) and access to content that they might not find offline.[17] Although this can create a positive user experience, it can also provide opportunities for extremism to spread.[17,46] While there is no typical profile of an extremist, risk factors that may make people more vulnerable to engaging with extremist content include lacking a sense of purpose or belonging, grievances with society, and experience of trauma.[17,47,48] Most researchers also agree that there is no typical process or pathway for forming extremist views.[49–52] The Internet may facilitate extremism in multiple ways, including recruitment, socialisation, communication, networking and mobilisation.[30,53]

Recruitment

Social media can enable groups to recruit members and reach new audiences. Recruitment strategies can make highly sophisticated use of technology to exploit existing grievances and cast doubt on mainstream society.[30,54] Groups may use offline trigger events (such as terror attacks, elections or financial crises) to promote extremist narratives.[18] Recruitment strategies may also appeal to a desire to belong, by presenting a positive image of life within an extremist group.[54] Many current and former extremists have expressed that their reasons for joining a group stem from anger or desire for a sense of unity. [47,55] Extremist groups may search social media to find vulnerable individuals and tailor their recruitment based on content that a user has posted.[56,57]

Socialisation

Socialisation refers to the process of internalising the norms of a group. Socialisation does not always involve ideological indoctrination. Extremist groups can provide a counter-culture with its own language and references that can be reinforced by a sense of collective struggle.[57,58] Socialisation may also make it harder for individuals to leave extremist groups.[30]

Individuals may develop a strong dedication to the online community and withdraw from offline peers, particularly if they are socially isolated. The absence of protective factors, such as a supportive network of friends and family,[59] means that online interactions can fulfil a sense of community,[32] and limit exposure to a narrow range of views.[60] Research suggests that

the Internet facilitates socialisation through 'echo chambers' (POSTnote 559) where extremist views are amplified and unchallenged.[30] Extremist groups also reward socialisation by using incentives (such as ranks and leader boards) to increase participation and loyalty.[29,30,61] Individuals may be motivated by this in order to gain status and infamy.[29]

The evidence for how the Internet contributes to socialisation is mixed.[62] It is more likely to occur through a combination of online and offline events and relationships. For example, offline trigger events (such as terror attacks) are correlated with an increase in extremist and hateful content online.[18] However, socialisation can occur entirely online, as the Internet can provide a social space for people that may otherwise be isolated from one another.[29,63]

Communication

Extremists use strategic communication to recruit members, raise funds, normalise extremist views, advise members on how to support the group, and gain publicity.[64–66] Organised groups often use smaller alt-tech platforms to co-ordinate mainstream campaigns, which occur on larger platforms or offline.[30] Propaganda may be circulated to a wider audience by including trending hashtags (searchable labels that refer to the topic of a social media post).[67–69] For example, a terrorist group used the hashtag #WorldCup when posting propaganda in 2014.[70] Extremists also use memes (typically easily shareable, supposedly amusing images containing symbolism recognisable to a subculture) to spread information to a chosen audience quickly.[71] During the 2019 General Election, extremist groups held meme-making competitions to have greater impact on online political discussion.[30]

Networking

Online platforms enable individuals to connect, irrespective of geographic location, and form global networks. Networking with prominent mainstream sympathisers provides extremists with a larger platform and can help to legitimise extreme views. Such groups may achieve greater recognition when public figures (such as politicians) share or reference their content online.[29,72] The Internet also enables collaboration in order to achieve a common goal. Extremist groups have worked together during elections to support or discredit a party or candidate.[29]

Mobilisation and attack

Online communication may enable extremist groups to co-ordinate members to undertake actions such as:

- Demonstrations and rallies that can intimidate the public and attract media attention. The impact of such events can be amplified by livestreaming them (broadcasting a live uncensored video in real-time on the Internet).[30]
- Cyber-attacks, such as hacking (gaining unauthorised access to a computer system by exploiting security weaknesses) and doxxing (releasing somebody's personal information online). These can intimidate society and draw media attention.[30]
- Contributing to terrorism, either through the coordination of an attack or through resources such as tutorials on making weapons. All five of the UK terrorist attacks in 2017 had an online element to them.[12] A survey of 227 convicted UK terrorists also revealed 44% had downloaded extremist media and at least 30% accessed extremist content online.[16]

Countering online extremism

Countering online extremism requires a coordinated approach across government, other public bodies, private companies and independent regulators, as well as the public.[12] However, this can be difficult to manage because views differ on where specific responsibilities lie for regulating the online environment (Box 2). The following sections consider the main options available for countering extremism. While technological and societal interventions are available, these responses are unlikely to eliminate extremism. People often experience difficulties in life prior to engaging in extremism, and reducing extremism may require intervening earlier to address these needs.[47] Therefore, researchers and charities suggest there should be a wider focus on offline and online safeguarding to make it harder for harmful activity to occur.[73,74]

Box 2: Regulating the online environment

Responsibility for regulating the Internet is shared across multiple public and private bodies, including:

- **Government departments**. The UK Government has responsibility for safeguarding UK citizens and interests from online harms. This duty is shared across multiple departments. For example, the Home Office is responsible for keeping the UK safe from the threat of terrorism and extremism, and the Department for Digital, Culture, Media and Sport for maintaining a safe and open Internet.[75,76]
- **Other public bodies**. If content breaches terrorism legislation, the Counter Terrorism Internet Referral Unit (based in the Metropolitan Police) refers content to host platforms for removal.[77] Since 2010, it has referred over 310,000 items of extremist content.[78]
- **Private companies**. Internet companies can remove extremist content based on breach of community guidelines.[79] There is cooperation between private companies. The Global Internet Forum to Counter Terrorism (GIFCT) was founded by Microsoft, Facebook, Twitter and YouTube. It maintains a database of known terrorist content that has been removed, preventing the same content being re-uploaded on any of the participating platforms.[80] In 2019, the GIFCT was criticised after it was unable to prevent over 800 versions of a terrorist attack video from appearing online.[81]
- **Regulators**. Although there is currently no regulator with specific responsibility for addressing online extremism, the 2019 Online Harms White Paper outlined proposals for the appointment of an independent regulator.[12]

Content removal

Extremist groups often have a coordinated strategy for rapid dissemination of content online.[82] Propaganda materials are typically uploaded across multiple platforms to increase longevity of exposure.[83] Online platforms are not currently obliged to remove extremist content in the UK if it does not fall under definitions of terrorism.[84] This is partially because of the issues around defining extremism. However, private companies may choose to remove content (Box 2).

Content posted online (such as terrorist videos) can be removed by human moderators or it can be detected and removed automatically. Automatic removal has many advantages, as methods cover more content in less time than human moderators. YouTube reports that 98% of violent extremism content that is removed is detected automatically,

and over half is removed within 2 hours.[85] Removing content soon after it is uploaded reduces the number of people exposed to it and helps to safeguard human moderators from upsetting material.[86] However, there are limitations to content removal:

- **Avoiding terms**. Users may deliberately avoid specific terms that may be flagged by automatic detection. Far-right groups may avoid fascist words and Islamist extremists may avoid ideology-specific words.[30] In 2016, online forum users began substituting certain terms with brand names to evade automatic detection.[87,88]
- **Images and videos**. Automatic detection often focuses on text but can be limited when faced with other forms of media. Images and videos only require slight modification in order to evade detection from a database of content that has been removed previously and is flagged for immediate removal if posted again (Box 2).[89]
- **Memes**. Memes often have indirect meanings, which may not be immediately interpretable to somebody outside of that community. This makes it hard for outsiders to judge what is harmful or extreme.[71]
- **False positives**. Automatic removal is prone to false positives, leading to issues around censorship if content is removed without justification. This can have consequences, such as censoring democratic debate, which may fuel extremist narratives around mistrust in the state.[84,88,90]
- **Lack of accountability**. Automatic detection may disproportionately impact certain groups, as datasets used for training algorithms contain biases. There is also a lack of transparency around how algorithms work, presenting issues for justifying content removal. [88,90]

Even with content removal, there are ways for extremists to disseminate content. It is easy to create 'throwaway' accounts to rapidly spread propaganda before content is removed.[91] Internet archiving services can also create links that allow content to remain accessible even when the original content is removed. Between 2015 and 2017, over 1 million archived links to Islamist extremist propaganda were found.[83,92] Also some platforms do not have a central server (POSTbrief 28) where all information is saved, making it challenging to remove content. People may also use browser tools that allow them to covertly comment on any mainstream website.[93] Additionally, extremists can operate on the dark web (POSTnote 488) where it is almost impossible to moderate or remove content.[94]

De-platforming

De-platforming is the removal of a group or individual from an online platform. In March 2018, Facebook banned a far-right group with 1.8 million followers. The group moved to the online forum Gab, which substantially reduced its following (44,000 as of March 2019).[95] De-platforming can be effective in reducing audience reach and preventing normalisation of harmful extreme views.[95] However, excluding users from mainstream sites can push them towards fringe platforms, where content is harder to detect and users are likely to be exposed to even more extreme views.[30] Extremist groups often aim to build mistrust in mainstream society and de-platforming reinforces narratives of oppressive censorship.[96] There is also the danger of disproportionately scrutinising certain groups, which can cause hostilities with those that feel targeted.[97]

Counter-narratives and alternative narratives

Counter-narratives directly oppose extremist messages. For example, extremist propaganda may advertise the positives of being in an extremist group. A counter-narrative would seek to expose the harmful reality of living under that group. An alternative narrative would highlight positive attributes outside of the extremist narrative, such as the benefits of living in a democratic society.[98,99] There is limited evidence on the effectiveness of these approaches, although counter-extremism organisations emphasise the importance of open dialogue for effective intervention.[59,100,101] Informal counter-narratives from trusted figures within the target community can be perceived as having higher credibility.[60] Researchers note that to be effective, counter-narratives may need to recognise the concerns of the extremist groups as genuine, which could be controversial if the concern is not widely seen as legitimate.[47]

Societal interventions

Many stakeholders believe that current counter-extremism responses are too focused on law and technology, and do not address the underlying reasons that people are drawn to extremist content.[28] Researchers and counter-extremism organisations highlight the importance of prevention before people have had the opportunity to engage in extremism, as well as rehabilitation.[47,102,103] Identifying underlying socio-economic and cultural contributors within society and implementing protective factors can reduce how many people develop extremist views.[47] The 2019 Integrated Communities Action Plan proposes using initiatives to promote values of tolerance and mutual respect online by empowering those who wish to challenge extremist views.[104] Proactive community training can be used to build digital literacy (POSTnote 608) so that users have the skills to assess the credibility of online content.[105,106] Stakeholders suggest that, to be effective, community interventions should be coordinated across multiple agencies (including police, social services and healthcare providers).[47]

Individual interventions

Individual interventions work with people who are deemed at risk of engaging with extremism. They can help support people to find more acceptable ways of expressing their feelings and meeting their needs.[48] However, not all people would be regarded (by themselves or others) as 'at risk' or vulnerable and may not want to engage with interventions.[48] Individuals may self-refer for intervention, be referred by members of the public, or be identified by police.[107] Individuals may by referred to the Prevent programme (run by the Home Office, counter-terrorism police and local community partnerships), which is currently under independent review.[108,109] Prevent aims to stop vulnerable people being drawn into terrorism. If somebody is considered at risk, they may be referred to the voluntary Channel programme where they receive mentorship and a support plan.[110] Between April 2018 and March 2019, 561 individuals received Channel support.[111] Stakeholders note that individuals could feel targeted by such programmes and that extremists may exploit this to undermine trust in the state.[112] Therefore, practitioners (including Prevent officers) advocate that such programmes should be delivered alongside wider community strategies to promote a safer society.[47,113]

POST is an office of both Houses of Parliament, charged with providing independent and balanced analysis of policy issues that have a basis in science and technology. POST is grateful to Clare Lally for researching this briefing, to the ESRC for funding her parliamentary fellowship, and to all contributors and reviewers. For further information on this subject, please contact the co-author, Dr Rowena Bermingham.

1. Brown, I. *et al.* (2015). Check the Web: Assessing the ethics and politics of policing the internet for extremist material. *VOX-Pol Netw. Excell.*
2. Manningham-Buller, E. (2005). The International Terrorist Threat and the Dilemmas in Countering It. *MI5 Security Service*.
3. HM Government *et al.* (2015). Counter-Extremism Strategy.
4. The Joint Committee on Human Rights (2016). The Proposed Counter-Extremism and Safeguarding Bill.
5. Commission for Countering Extremism (2019). Challenging Hateful Extremism.
6. Commission for Countering Extremism (2020). Police Recall Guidance That Includes Extinction Rebellion. *Extremism Commission Blog*.
7. McCann, C. (2019). The UK's "Hateful Extremism" report: more muddled thinking on an already over-crowded pitch. *openDemocracy*.
8. Lowe, D. (2017). Prevent Strategies: The Problems Associated in Defining Extremism. *Stud. Confl. Terror.*, Vol 40, 917–933.
9. HM Government (2006). Terrorism Act 2006.
10. Berger, J. M. (2018). Extremism. The MIT Press Essential Knowledge.
11. Home Office (2019). Revised Prevent duty guidance: for England and Wales.
12. HM Government *et al.* (2019). Online Harms White Paper.
13. Meleagrou-Hitchens, A. *et al.* (2017). Research Perspectives on Online Radicalisation. *VOX-Pol Netw. Excell.*
14. Ganesh, B. *et al.* (2020). Extreme Digital Speech: Contexts, Responses, and Solutions. *VOX-Pol Netw. Excell.*
15. Hassan, G. *et al.* (2018). Exposure to Extremist Online Content Could Lead to Violent Radicalization: A Systematic Review of Empirical Evidence. *Int. J. Dev. Sci.*, Vol 12, 71–88.
16. Gill, P. *et al.* (2017). Terrorist Use of the Internet by the Numbers. *Criminol. Public Policy*, Vol 16, 99–117.
17. Pasculli, L. (2020). The Global Causes of Cybercrime and State Responsibilities. Towards an Integrated Interdisciplinary Theory. *J. Ethics Leg. Technol.*,
18. Awan, I. *et al.* (2019). Extremism Online: Analysis of Extremist Material on Social Media.
19. Ahmed, M. (2020). After Christchurch: How Policymakers Can Respond to Online Extremism. *Institute for Global Change*.
20. Piazza, J. A. *et al.* (2019). The Online Caliphate: Internet Usage and ISIS Support in the Arab World. *Terror. Polit. Violence*.
21. Harrison, S. (2018). Evolving Tech, Evolving Terror. *Center for Strategic and International Studies*.
22. Corera, G. (2014). Police Play Cat and Mouse with Online Jihadists. *BBC News*.
23. Tech Against Terrorism (2019). Insights from the Centre for Analysis of the Radical Right's Inaugural Conference in London. *Tech Against Terrorism*.
24. Twitter (2019). Transparency Report - June 2019. *Twitter*.
25. Yang, K.-C. *et al.* (2019). Arming the public with artificial intelligence to counter social bots. *Hum. Behav. Emerg. Technol.*, Vol 1, 48–61.
26. Ribeiro, M. H. *et al.* (2019). Auditing Radicalization Pathways on YouTube. *arXiv*.
27. Reis, J. *et al.* (2015). Breaking the News: First Impressions Matter on Online News. *Proc. Ninth Int. AAAI Conf. Web Soc. Media*, 357–366.
28. Glazzard, A. (2019). Shooting the Messenger: Do Not Blame the Internet for Terrorism. *R. United Serv. Inst.*, Vol 39.
29. Nagle, A. (2017). Kill All Normies: Online Culture Wars from 4chan and Tumblr to Trump and the Alt-Right. Zero Books.
30. Ebner, J. (2020). Going Dark: The Secret Social Lives of Extremists. Bloomsbury Publishing.
31. Europol (2019). Europol and Telegram Take On Terrorist Propaganda Online. *Europol*.
32. de Koster, W. *et al.* (2008). 'Stormfront Is Like a Second Home to Me'. *Inf. Commun. Soc.*, Vol 11, 1155–1176.
33. Bowman-Grieve, L. (2009). Exploring "Stormfront": A Virtual Community of the Radical Right. *Stud. Confl. Terror.*, Vol 32, 989–1007.
34. Clemmow, C. *et al.* (2020). The Base Rate Study: Developing Base Rates for Risk Factors and Indicators for Engagement in Violent Extremism. *J. Forensic Sci.*
35. Borum, R. (2011). Radicalization into Violent Extremism I: A Review of Social Science Theories. *J. Strateg. Secur.*, Vol 4, 7–36.
36. Reeve, Z. (2019). Engaging with Online Extremist Material: Experimental Evidence. *Terror. Polit. Violence*, 1–26.
37. Hansford, S. (2020). A National Asset: Why Data is Vital for UK Citizens. *Computer Business Review*.
38. James, M. (2020). Digital sovereignty: Time for Britain to Take Back Control of our Data. *CityAM*.
39. van Mill, D. (2018). Freedom of Speech. in *The Stanford Encyclopedia of Philosophy*. (ed. Zalta, E. N.) Metaphysics Research Lab, Stanford University.
40. Huszti-Orban, K. *et al.* (2020). The Impact of "Soft Law" and Informal Standard-Setting in the Area of Counter-Terrorism on Civil Society and Civic Space.
41. Oltermann, P. (2018). Tough new German law puts tech firms and free speech in spotlight. *The Guardian*.
42. Tworek, H. *et al.* (2019). An Analysis of Germany's NetzDG Law. *Transatlantic Work Group*.
43. Echikson, W. *et al.* (2018). Germany's NetzDG: A Key Test for Combatting Online Hate. *Cent. Eur. Policy Stud.*,
44. Human Rights Watch (2018). Germany: Flawed Social Media Law. *Human Rights Watch*.
45. DeNardis, L. *et al.* (2015). Internet governance by social media platforms. *Telecommun. Policy*, Vol 39, 761–770.
46. Babbs, D. (2019). Academic Research About Online Disinhibition, Anonymity, and Online Harms. *Clean Up the Internet*.
47. Bellis, M. A. *et al.* (2019). Preventing Violent Extremism in the UK: Public Health Solutions. 1–84. Public Health Wales.
48. Logan, C. *et al.* (2019). Violent Extremism: A Comparison of Approaches to Assessing and Managing Risk. *Leg. Criminol. Psychol.*, Vol 24, 141–161.
49. Bjørgo, T. (2011). Dreams and Disillusionment: Engagement in and Disengagement from Militant Extremist Groups. *Crime Law Soc. Change*, Vol 55, 277–285.
50. Bartlett, J. *et al.* (2011). The Edge of Violence: Towards Telling the Difference Between Violent and Non-Violent Radicalization. *Terror. Polit. Violence*, Vol 24, 1–21.
51. Post, J. M. (2015). Terrorism and Right-Wing Extremism: The Changing Face of Terrorism and Political Violence in the 21st Century: The Virtual Community of Hatred. *Int. J. Group Psychother.*, Vol 65, 242–271.
52. Richards, A. (2015). From Terrorism to 'Radicalization' to 'Extremism': Counterterrorism Imperative or Loss of Focus? *R. Inst. Int. Aff.*
53. Benigni, M. C. *et al.* (2017). Online Extremism and the Communities that Sustain It: Detecting the ISIS Supporting Sommunity on Twitter. *PLOS ONE*.
54. Speckhard, A. *et al.* (2015). Eyewitness Accounts from Recent Defectors from Islamic State: Why They Joined, What They Saw, Why They Quit. *Perspect. Terror.*, Vol 9, 95–117.
55. Allan, H. *et al.* (2015). Drivers of Violent Extremism: Hypotheses and Literature Review. *R. United Serv. Inst.*

56. Scanlon, J. R. *et al.* (2014). Automatic Detection of Cyber-Recruitment by Violent Extremists. *Secur. Inform.*, Vol 3.
57. Jost, J. T. *et al.* (2007). Shared Reality, System Justification, and the Relational Basis of Ideological Beliefs. *Soc. Personal. Psychol. Compass*, Vol 2, 171–186.
58. Whitehouse, H. *et al.* (2017). The Evolution of Extreme Cooperation via Shared Dysphoric Experiences. *Sci. Rep.*, Vol 7, 1–10.
59. Beutel, A. *et al.* (2016). Guiding Principles for Countering and Displacing Extremist Narratives. *Cent. Study Terror. Polit. Violence J. Terror. Res.*, Vol 7.
60. Lee, B. (2020). Countering Violent Extremism Online: The Experiences of Informal Counter Messaging Actors. *Policy Internet*, Vol 12, 66–87.
61. Schlegel, L. (2018). Terrorism and the Gamification of Radicalization. *The Defense Post*.
62. O'Hara, K. *et al.* (2015). Echo Chambers and Online Radicalism: Assessing the Internet's Complicity in Violent Extremism. *Policy Internet*, Vol 7, 401–422.
63. Moonshot CVE (2019). It's Time to Save Incels From Themselves. *Moonshot CVE*.
64. Winter, C. (2019). Making Sense of Jihadi Stratcom: The Case of the Islamic State. Vol 13, 54–62.
65. Bockstette, C. (2008). Jihadist Terrorist Use of Strategic Communication Management Techniques. Defense Technical Information Center.
66. Williams, L. (2016). Islamic State Propaganda and the Mainstream Media. *Lowy Institute*.
67. Girginova, K. (2017). Hijacking Heads & Hashtags. *Glob.-E*, Vol 10.
68. Farwell, J. P. (2014). The Media Strategy of ISIS. *Surviv. Glob. Polit. Strategy*, Vol 56, 49–55.
69. Alfifi, M. *et al.* (2018). Measuring the Impact of ISIS Social Media Strategy. *Stanf. Netw. Anal. Proj.*
70. AFP (2014). How ISIS Used Twitter and the World Cup to Spread Its Terror. *The Telegraph*.
71. Lee, B. (2020). 'Neo-Nazis Have Stolen Our Memes': Making Sense of Extreme Memes. in *Digital Extremisms: Readings in Violence, Radicalisation and Extremism in the Online Space*. Palgrave Studies in Cybercrime and Cybersecurity.
72. BBC News (2017). Donald Trump Retweets Far-Right Videos. *BBC News*.
73. Babbs, D. (2020). New Year, New Internet? Why it's Time to Rethink Anonymity on Social Media. *openDemocracy*.
74. Krasodomski-Jones, A. (2019). The government is attempting to fix the internet - and we should be cautiously optimistic. *New Statesman*.
75. Department for Digital, Culture, Media and Sport (2019). Annual Report and Accounts for the Year Ended 31 March 2019.
76. Home Office (2020). About Us. *Home Office Website*.
77. HM Government (2020). Report online material promoting terrorism or extremism. *Home Office Website*.
78. Counter Terrorism Policing (2018). Together We're Tackling Online Terrorism. *Counter Terrorism Policing*.
79. Home Affairs Commitee (2017). Hate Crime: Abuse, Hate and Extremism Online. *House of Commons*.
80. Global Internet Forum for Countering Terrorism (2020). GIFCT Website. *Global Internet Forum for Countering Terrorism*.
81. Hern, A. (2019). Facebook and YouTube defend response to Christchurch videos. *The Guardian*.
82. HM Government (2019). Online Harms White Paper.
83. Flashpoint (2018). An Analysis of Islamic State Propaganda Distribution.
84. van der Vegt, I. *et al.* Shedding Light on Terrorist and Extremist Content Removal. *R. United Serv. Inst.*
85. Wojcicki, S. (2017). Expanding Our Work Against Abuse of Our Platform. *Official YouTube Blog*.
86. BBC News (2020). Social Media Moderators Asked to Sign PTSD Forms. *BBC News*.
87. Hagen, S. (2017). Alt-Right Open Intelligence Initiative. *Digital Methods Inititiative*.
88. Vidgen, B. *et al.* (2019). Challenges and Frontiers in Abusive Content Detection. *Proc. Third Workshop Abus. Lang. Online*, 80–93.
89. Ghosh, S. (2019). Tech Firms are Scrambling to Keep Dozens of Copies of Videos Showing the New Zealand Mosque Shootings off their Platforms. *Business Insider*.
90. Ofcom (2019). Use of AI in Online Content Moderation.
91. Macdonald, S. *et al.* (2019). Daesh, Twitter and the Social Media Ecosystem: A Study of Outlinks Contained in Tweets Mentioning *Rumiyah*. *RUSI J.*, Vol 164, 60–72.
92. Kelion, L. (2018). IS Propaganda 'Hidden on Internet Archive'. *BBC News*.
93. Dissenter (2020). Dissenter: Free Speech Web Browser. *Dissenter*.
94. Weimann, G. (2016). Going Dark: Terrorism on the Dark Web. *Stud. Confl. Terror.*, Vol 39, 195–206.
95. Nouri, L. *et al.* (2019). Following the Whack-a-Mole: Britain First's Visual Strategy from Facebook to Gab. *R. United Serv. Inst.*
96. Hope Not Hate (2018). Alt-Tech: Far-Right Safe Spaces Online. *Hope Not Hate*.
97. Ebner, J. The Rage: The Vicious Circle of Islamist and Far-Right Extremism. Bloomsbury Publishing.
98. Integrity UK (2014). Telling Compelling Stories. *Integrity UK*.
99. YouTube (2020). Creators for Change. *YouTube*.
100. Ferguson, D. K. (2016). Countering violent extremism through media and communication strategies. *Norwich Partnersh. Confl. Crime Secur. Res. Univ. East Angl.*
101. Ramalingam, V. (2019). Why I Talk to White Supremacists. *TEDxLondonWomen*.
102. UNESCO (2016). Preventing Violent Extremism. *United Nations Educational, Scientific and Cultural Organization*.
103. Harris-Hogan, S. *et al.* (2016). What is Countering Violent Extremism? Exploring CVE Policy and Practice in Australia. *Behav. Sci. Terror. Polit. Aggress.*, Vol 8, 6–24.
104. Ministry of Housing, Communities & Local Government (2019). Integrated Communities Action Plan.
105. JAN Trust (2018). Web GuardiansTM. *JAN Trust*.
106. Me & You Education (2020). Our Training. *Me & You Education*.
107. National Police Chiefs' Council (2020). Prevent. *NPCC*.
108. UK Counter-Terrorism Policing (2018). Prevent: What We Do. *Counter Terrorism Policing*.
109. Home Office (2019). Independent Review of Prevent: Terms of Reference.
110. Home Office (2019). Channel Guidance.
111. Home Office (2019). Individuals referred to and supported through the Prevent programme: England and Wales, April 2018 to March 2019.
112. Choudhury, T. *et al.* (2011). The Impact of Counter-Terrorism Measures on Muslim Communities. *Equal. Hum. Rights Comm.*, 151–181.
113. Prevent, Metropolitan Police, Personal Communication.