To be given at Swansea University, Terrorism and Social Media Conference (28 June 2022)

"We are assured that the world is getting more and more united and growing into a brotherly community by the reduction of distances and the transmission of ideas through the air. Alas, put no faith in such a union of peoples."

(*The Brothers Karamazov*, trans. Magarshack)

"The Matrix is everywhere. It is all around us."

(*The Matrix*, dir. the Wachowskis)

RIGHTS AND VALUES IN COUNTER-TERRORISM ONLINE

Summary

It is easy to say that counter-terrorism activity online should be guided by reference to fundamental human rights. However, identifying who holds those rights, who enforces those rights, and who has a duty to respect those rights is far from straightforward in the online context. Understanding when qualification of those rights may be justified in the public interest is also close to impossible given the impenetrable complexity of the internet. I argue that values are more coherent than rights and that the principal online value is freedom of expression. I draw 5 tentative conclusions on how the value of freedom of expression can be sufficiently protected, whilst recognising that some content moderation is justified in the interest of preventing real world violence.

No Surprise That Terrorists Use the Internet

1.1. It was never going to be long before the terrorist worm entered the online paradise. As much as terrorists exploit air by breathing, they exploit the internet by typing, tapping and scrolling like the rest of us.

1.2. Listing all the ways in which terrorists and their sympathisers use the internet would be both quaint and impossible. A vast literature exists to illustrate this at general[1] and highly granular levels[2]. Suffice it to say that they have proven to be early adopters at exploiting all the different uses (termed 'affordances') to which online platforms lend themselves[3].

[1] E.g. UN Office on Drugs and Crime, 'The Use of the Internet for Terrorist Purposes' (New York, 2012).
[2] E.g. Macdonald, S., Rees, C., S., J., 'Remove, Impede, Disrupt, Redirect: Understanding and Combating Pro-Islamic State Use of File-Sharing Platforms' (Resolve Network, April 2022). Tech Against Terrorism identify the following categories of platforms used to disseminate content: beacons, content stores, aggregators and circumvention platforms.
[3] For an accessible history of evolving use of platforms between 2003-2019, see Williams, H., Evans, A., Ryan, J., Mueller, E., Downing, B., 'The Online Extremist Ecosystem' (Rand Corporation, December 2021).

1.3. These **agile, adaptive and savvy tech performers**[4] are alert to endeavours to remove content or shut down channels, resulting in a frequent exodus from larger to smaller platforms[5]. The rate at which the Christchurch attack video was uploaded in the immediate aftermath, including in files that had been deliberately altered to frustrate blocking technologies[6], suggest that there are tens if not hundreds of thousands of individuals who know that the authorities worry about the presence of that sort of material, but – whether for free speech or more sinister motives – are determined to keep it in circulation.

1.4. The harm feared is real world terrorist violence. I will explain later in this paper why we need to be clear about the harm that counter-terrorism seeks to address. The question is what steps should be taken by way of countering this feared harm. In general, this involves tech companies:

- Stopping content being uploaded to platforms.
- Taking down content that has already been uploaded.
- Closing channels.
- Changing algorithmic suggestions which might promote terrorist content.
- Disrupting the ability to share terrorist content widely by changing affordances on platforms.
- Changing the context in which material appears (for example, having counterpoint narratives appearing alongside potential terrorist content).

Fundamental Rights

1.5. At a superficial level, the **interpenetration of the online and offline worlds** means that the interests protected by fundamental rights are no less relevant for the internet than *i(n) r(eal) l(ife)*. Terrorist attacks are planned and facilitated online, and governments that vacated the field entirely would fail to honour their basic obligation to keep citizens safe.

1.6. Yet an analysis based on enforceable rights of citizens *against the state* is inadequate for identifying a principled response to online risks. Whilst fundamental rights have frequently been invoked to require governments to put in place protection against harm from non-state actors[7], the key online actors are not states but tech companies and the billions of user-generators scattered across the globe. As to their relative power:

- The nations of the world have failed to identify and impose a common definition of terrorism, let alone come up with an enforcement model for

[4] Rasmussnen, N., GIFCT Executive Director, 'The Dynamic Terrorism Landscape and What it Means for America' (written testimony to US House of Representatives Commission on Homeland Security, 2.2.22).
[5] HM Government, 'Impact Assessment Online Safety Bill' (311.22) at para 357.
[6] New Zealand Government, '2021: Digital Violent Extremism Transparency Report' (2022), p31.
[7] As in KU v Finland App.No. 2872/02 (2.12.08) in which the ECtHR held that the government of Finland needed to have protective laws against online sexual abuse.

states to control the behaviour of irresponsible users or platforms in overseas jurisdictions.

- The complexity and potency of networked computers means that it is not governments but platforms, or perhaps just a handful of brilliant technicians, who dictate what is feasible. For example, if platforms choose end-to-end encryption or encryption of metadata that will limit the ability of state agencies to scan for content.
- For example, in a case brought against Italy by a parents' association whose children had been targeted by obscene spam, the European Court of Human Rights agreed that the recipients' private life had been interfered with. However, the application was inadmissible because there was little that Italy could have done by way of counter measures[8].

1.7. Fundamental rights connote corresponding duties requirements of due process. The ultimate expression of due process would be a platform which, in the interests of safeguarding rights, refused to remove any content without a court order. Indeed, it has been suggested that independent adjudication is a necessary component of any moderation of 'illegal content' [9].

1.8. However, as Evelyn Douek has persuasively argued, a traditional rights analysis of individual versus state implies notions of procedural protection that are alien to this environment[10]. Because of the scale of content that may need to be processed, the intervention of a human moderator, subject to judicial review before an independent tribunal, cannot possibly be guaranteed for every takedown decision made by a platform – even disregarding the jurisdictional difficulties of identifying a moderator and judge who could authoritatively decide on content posted anywhere in the world.

1.9. Moreover, billions of individual users and tech companies themselves[11] are themselves the recipients of rights. Most significantly, the United States Constitution's **First Amendment** gives overriding primacy to freedom of expression even in cases where expression amounts to calls to violence and criminality[12].

1.10. Perhaps just as significant, the internet has established **social expectations**[13] whose reversal is now inconceivable. These expectations must be considered when identifying and ranking rights or values. Measures to limit online freedoms in the

[8] Muscio v Italy App.No. 31358/03 (13.11.07).
[9] Smith, G., 'Should We Be Building Online Prior Restraint Machines' (Society for Computers and Law, 22.1.18).
[10] Douek, E., Content Moderation as Administration (January 10, 2022). forthcoming Harvard Law Review Vol. 136
[11] Cf. Case of Markt Intern Verlag GMBH and Klaus Beermann v Germany, App.No.10572/83 (20.11.89); Citizens United v Federal Election Committee, 558 U.S. 310 (2010).
[12] Save in cases of "imminent lawless action": Brandenburg v Ohio, 395 U.S. 444 (1969). The position under UK common law and the ECHR is of course quite different because the right or freedom of expression may be proportionately curtailed in the wider public interest, specifically, in the terrorist context, in the interests of national security. Conversely, the UK (as a result of the ECHR) has adopted protections for privacy that go far beyond those applicable in the US. The position of the platform Gab is to enable any content that is protected by the First Amendment: Annual Report, 22 May 2020.
[13] Cengiz and Others v Turkey, App.No. 48226/10 and 14027/11 (12.12.15) at paras 49 and 52.

name of counter-terrorism that failed to accommodate demand for instantaneous information access and exchange, and faster and ever more efficient services, would not be tolerated in an open society[14].

1.11. The public backlash against **OnlyFans'** decision to ban sexually explicit material on child safety grounds, forcing a reversal within 6 days[15], and consumer demand for the most secure levels of **encryption** despite the grave risk of consequence-free exploitation by to terrorists and child sex abusers[16], illustrate the power of the market and however imperfectly, social expectations. Moreover, democratic states such as the United Kingdom have hitched themselves to powerful producer interests in accepting a free internet as vital for **driving economic growth and providing innovative solutions**[17].

1.12. Added to this, understanding the **trade-offs** between counter-terrorism and internet functionality is a closed book to those without detailed insider technical knowledge. This is relevant to how lawyers consider whether a qualification of fundamental rights in the wider public interest is justified[18].

- It is difficult for the public or policy-makers to evaluate the argument that regulatory burden would be a terminal threat to start-ups, or that practical content moderation is only possible through use of algorithms or machine learning that would have unintended consequences[19]. There is a powerful case for greater transparency from tech companies to inform this debate[20].
- Governments are rightly wary of solidifying gains made by powerful producers – the Facebooks, Googles and Amazons – and recognise that a vibrant internet economy must embrace challengers. For example, the UK-based platform BitChute was established by 2 individuals in 2017 and by 2022 it had 12 employees and tens of millions of monthly visits[21]. Unfortunately, this free speech platform[22] quickly became a vehicle for **Neo-Nazi propaganda**[23]. Accommodating the business model of small platforms necessarily limits the extent to which regulation can be imposed.
- There is also a more profound anxiety relating to those services on which the architecture of the internet depends (such as domain name providers), about

[14] For disabled users, the freedoms and opportunities created by the internet may be far more important than these advantages.
[15] Columbo, C., 'The history of OnlyFans: How the controversial platform found success and changed online sex work', Insider (14.9.21).
[16] Buhler, K., 'The Rising Consumer Demand for Data Privacy and Autonomy', Sequoia (18.11.21).
[17] Declaration for the Future of the Internet (April 2022) to which the UK, US and EU Member States among others are signatories.
[18] The proportionality exercise requires consideration of a 'fair balance' between individual rights and public interests: Bank Mellat v HM Treasury (No 2) [2013] UKSC 39 at para 20, Lord Sumption.
[19] Gillespie, T., et al, 'Expanding the debate about content moderation: scholarly research agendas for the coming policy debates', Internet Policy Review Vol.9 Issue 4 (21 October 2020).
[20] Douek, E., supra.
[21] Bitchute, Transparency Report (June 2022).
[22] Trujillo, M, Gruppi, M., Buntain, C, Horne, B., 'What is BitChute? Characterizing the "Free Speech" Alternative to YouTube' (31st ACM Conference on Hypertext and Social Media, July 13–15, 2020).
[23] 'Hate Fuel: the online world fuelling far right terror', (CST, 1 May 2020).

the desirability of imposing rules at all for fear of politicising the internet leading to its eventual fragmentation[24].

1.13. Finally, the common practice of referring to human rights **collectively** begs the question of what precise rights are being invoked. The Santa Clara Principles refer to "human rights" as a single block without distinction, save in a passage which refers "particularly [to] the rights to freedom of expression and non-discrimination"[25]. The implication is that freedom of expression is paramount (noting that non-discrimination is really a principle rather than a right).

1.14. This leads to serious ambiguity. It appears that all rights are in play – including the right to life and bodily integrity – but that freedom of expression is most deserving of protection. In some contexts it would be possible to conclude that giving priority to freedom of expression as a fundamental human right provides cover for perpetuating profitable practices on the part of tech companies who are blind to harms and antagonistic to any restrictions. In similar vein, references to legal clarity and parsimoniousness when interfering with fundamental rights are fine principles but they also result in fewer and cheaper rules for tech companies to implement.

1.15. So it is unsurprising that, as the lawyer and author Graham Smith puts it, rights have come to mean different things to different people and appeals to fundamental values have come to resemble "policy advocacy clothed in the language of rights"[26].

Values

1.16. Identifying rights, with the corollary that another party has a duty to respect those rights, is a tough analytical gig. It is conceptually easier to avoid the language of fundamental rights and refer instead to **values** in the sense of the priorities which should guide online counter-terrorism.

1.17. For example, although Tech Against Terrorism's states that its "…aim is to counter terrorist use of the internet whilst respecting human rights"[27], this surely does not refer a duty on the part of Tech Against Terrorism to protect the human rights of unspecified rights-holders. It is more coherent to understand this mission statement as a commitment to encouraging governments and tech companies to recognise certain values in the decisions they make.

1.18. Similarly, although the second iteration of the Santa Clara Principles was designed to "support companies to comply with their responsibilities to respect human rights", it is telling that the principles themselves refer to "human rights

[24] Bennett, A., Garson, M., Boakye, B., Beverton-Palmer, M., Erzse, A., 'The Open Internet on the Brink: A Model to Save Its Future' (Tony Blair Institute for Global Change, 2021).
[25] Santa Clara Principles 2.0, Foundational Principles, Chapeau and para 1.
[26] 'Speech vs. Speech', (www.cyberleagle.com, 22 June 2021).
[27] TCAP Transparency Report (March 2022) at para 4.2.1.

considerations"[28] (my emphasis) - a tacit recognition that human rights in the reciprocal-duty sense do not apply.

1.19. So what are the values that ought to have priority in counter-terrorism? Before turning to freedom of expression, I suggest that this paper has already identified one that does not derive expressly from a traditional list of human rights[29] but which is obviously supportive of the enjoyment and ranking of values in this context – that is the value of transparency.

Freedom of Expression and Association

1.20. I start with value of free expression because, applying an evidence-based approach, it is *certain* that any counter-terrorism measure online will remove or limit people's ability to read see or hear the words, images and sounds they would otherwise wish to encounter; to express the words, images and sounds that they would otherwise like to express; or to engage in the type of online associations with other individuals that that they would otherwise like to engage in.

1.21. Freedom of expression encompasses the ability to impart and receive information and, as well as being intrinsically important, is considered to protect three values: those of truth, democracy and individual autonomy or self-fulfilment[30]. It has long been formulated in terms of receiving and imparting information "regardless of frontiers"[31].

1.22. The first of the three values, **truth**, is particularly resonant in the terrorism context. Terrorist attacks change nations. Activities of violent diaspora groups linked to overseas conflicts, or violent domestic movements, are part of world history and personal experiences. The terrorist/ freedom fighter dilemma is inescapable, and content posted for sinister reasons may nonetheless be a true record. Where content, however disturbing, is used to tell the truth about an individual's own experience, the law rightly regards the ability to do this as a "basic right" to which the law gives "a very high level of protection"[32]: in less legalistic terms, truth is "our richest merchandise"[33].

1.23. Since terrorist groups (such as Da'esh/ Islamic State) are principal actors in world-changing events, there is truth value in knowing or establishing the truth about

[28] Ibid, Foundational Principles, para 1.
[29] Such as the European Convention on Human Rights.
[30] These three values identified in Frederich Shauer, 'Free Speech: A Philosophical Enquiry' (Cambridge, 1982) were deployed by Lord Steyn in R v Secretary of State for the Home Department, *ex parte Simms* [2000] 2 AC 115 at 126.
[31] Article 19 Universal Declaration of Human Rights; Article 19 International Covenant on Civil and Political Rights; Article 10 ECHR.
[32] James Rhodes (Appellant) v OPO (by his litigation friend BHM) and another (Respondents) [2015] UKSC 32 at para 76, 77.
[33] Milton, J., 'Areopagitica' (1664).

groups or individuals pursuing social change through violence[34]. This places a value on access to **disturbing footage and blatant propaganda** – and not merely for academics or journalists.

- Its value to the historical record may not be obvious at the time.
- The purpose for which the information was posted online does exclude its utility in establishing the truth.
- The value of compiling a truthful record provides a strong imperative to allow content to be posted and once posted to secure it, so that all internet content is kept for future reference and not destroyed.

1.24. Whilst online content is not notorious for its adherence to truth, even demonstrably false content will generate true metadata: a time and date, technical information, and potentially clues as to the identity of the person who posted the falsehood. Truth may also be expressed in different guises (by novels or songs[35], such as **nasheeds or anashids,** a subset of which celebrates jihadi violence and are produced by IS/Da'esh[36]) and freedom of expression protects choice as to how to express truth[37] including offensively[38]. Content exposes how individuals were using the internet at that point in time: referred to as the 'use meaning' rather than the 'representational meaning' of the words, images and sounds encountered[39]. This is consistent with the position taken by search engines such as Google to index, and ultimately make available to the general user, all surface web content[40].

1.25. It follows that there is value in protecting content even if the motives of the content provider are so abusive that they themselves may be said to have forfeited reliance on a fundamental right[41].

---

[34] Which is why, as Professor Maura Conway points out, removing only violent propaganda made by terrorist organisations and leaving up the happy material (pictures of nurseries etc.) distorts the truth about the nature of these organisations.

[35] Article 10 ECHR applied as much to the songs of Pussy Riot as to the symbolic display of dirty laundry near the Hungarian Parliament: *Mariya Alekhina and Others v Russia* (2019) 68 EHRR 14.

[36] Velasco-Pufleau, L., 'Jihadi Anashid, Islamic State Warfare and the Agency of Sound', Crime and Music, Springer, pp.233-243, 2021.

[37] Per Lord Neuberger in Rhodes, supra, citing Campbell v MGN Ltd [2004] UKHL 22, [2004] 2 AC 457, para 59, and In re Guardian News and Media Ltd [2010] UKSC 1, [2010] 2 AC 697, para 63.

[38] Sedley LJ in *Redmond-Bate v Director of Public Prosecutions* (1999) 7 BHRC 375, [20].

[39] Blocher, J., 'Nonsense and the Freedom of Speech: What Meaning Means for the First Amendment', [2014] Duke Law Journal vol.63: 1423.

[40] Google, 'Maximise access to information' https://www.google.com/search/howsearchworks/mission/open-web/ accessed 13.5.22).

[41] The ECtHR held in *Norwood v United Kingdom* App No 23131/03 at para 4 that a poster advocating the removal of Islam from the UK because of the 9/11 attacks did not enjoy the protection of Article 10 in light of Article 17 (abuse of rights). The invocation of Article 17 in this context looks like an overreaction and is not without its critics: A Buyse, "Dangerous Expressions: The ECHR, Violence and Free Speech" (2014) 63(2) *International & Comparative Law Quarterly 491*. The better explanation may well be that Mr Norwood's conviction for incitement to hatred and violence was justifiable because his freedom of expression was acceptably qualified in the wider public interest.

1.26. The mission of sites such as the Internet Archive (archive.org) is not just to preserve but to maintain general availability: "Universal access to All Knowledge"[42]. Europol assessed that **jihadi propogandists were exploiting the Internet Archive** for their own purposes[43] by making use of its permanency and openness.

1.27. Next, the ability of individuals to participate in public decision-making, and therefore **democracy**, is nothing without freedom of speech: the free flow of information and ideas informs political debate and voting, is a safety valve because people are more willing to accept adverse decisions if they can seek to influence them through, and acts as a brake on the abuse of power by public officials and others[44].

1.28. The internet is one of, if not the principal means[45] by which individuals find information relevant to public life, whether through traditional newspapers and broadcasters that have gravitated online, or through untrained members of the public operating as citizen journalists such as *The Sandwell Skidder*[46] or simply users of social media[47].

1.29. Images from conflict areas in which terrorists are active, including of the terrorist acts of the **so-called Islamic State Beatles in Syria**, are important documents in informing democratic debate on vital matters of public policy. Sometimes editorial judgment may call for the use of shocking images including what might be described as terrorist propaganda.

1.30. The third interest instrumentally protected by freedom of expression is **individual autonomy or self-fulfilment**.

- For some people, online engagement will be vital to the promotion of these interests: for example, video-conferencing by someone permanently confined to bed, or membership of an online support group for sufferers from an extremely rare and disabling disease. Freedom of expression underpins freedom to associate[48] to which digital technology and online spaces are now integral[49].
- It is true that much online engagement is objectively deleterious to personal development. Individuals do not bring their 'best selves' to the internet, as David Baddiel has memorably illustrated[50]. The combination of wanting a

[42] "About the Internet Archive" (archive.org/about/, accessed 11 May 2022).
[43] Europol, 'Jihadist content targeted on Internet Archive platform' (press release, 16 July 2021).
[44] Lord Steyn, ex parte Simms, supra; *R v Shayler* [2003] 1 AC 247, [21]. In an era of disinformation it would be naïve not to recognise that the internet calls into question JS Mill's characterisation of the free competition of ideas as the best way to separate falsehoods from fact.
[45] Cf. Mustafa v Sweden, 16.12.08 in which the internet was the only means of hearing news in the applicant's home language.
[46] McNally v Saunders [2021] EWHC 2012 (QB), with thanks to Graham Smith for this reference.
[47] *Magyar Helsinki Bizottság v Hungary* (2020) 71 EHRR 2 at para 168.
[48] Article 20 Universal Declaration of Human Rights; Articles 21 and 22 International Covenant on Civil and Political Rights; Article 11 ECHR.
[49] UN Special Rapporteur on the rights to freedom of peaceful assembly and of association, Report to Human Rights Council (17 May 2019).
[50] David Baddiel: Social Media, Anger and Us (BBC 2, 14 December 2021).

tribe, vying for shock value, and the freedom from convention that comes with (generally) anonymous engagement means that online personalities may be considerably more sympathetic to terrorist violence than their owners are away from the screen – indeed different personalities all together[51].

- But, despite the occasional desire of autocratic governments and frustrated parents to pull the plug on the internet, it is now too central to the way we all communicate and find meaning to wish it away.

Privacy and Correspondence

1.31. Privacy[52] protects expression of individuality and an inner life, the facilitation of trust, friendship and intimacy, the securing of other rights (for example by protecting journalistic sources), and empowering individuals against the state[53]. These interests will be in play when a lonely individual, perhaps a neurodivergent adolescent with no friends at school finds purpose and solace through membership of an **online group**[54] **of Second World War enthusiasts**: the problem comes when members of the group start to fixate on Nazi memorabilia, then violence against Jews and Muslims.

1.32. Protection of privacy is often cast as a value or right in opposition to freedom of expression. In this jurisdiction, the law may require limits to be placed on access to personal information, including information previously placed voluntarily in the public domain because personal information may be leaked, or aggregated, or analysed, or stored in a way that inhibits personal life. Privacy rights are invoked against state surveillance of personal communications email[55] or bulk data[56].

1.33. However, as the ability of the state to monitor content effectively is degraded by sheer volume, technical measures such as end-to-end encryption and DNS over HTTPS, and the difficulty of attributing content to users, public authorities cannot be relied upon to identify terrorist content online and deal with its human

[51] Blumer, T., Döring, N., "Are we the same online? The expression of the five factor personality traits on the computer and the Internet", Cybersecurity 6(3) (December 2012).
[52] Article 12 Universal Declaration of Human Rights; Article 17 International Covenant on Civil and Political Rights; Article 8 ECHR.
[53] Anderson, D., 'A Question of Trust: Report of the Investigatory Powers Review' (June 2015) at paras 2.10 to 2.13.
[54] Separating out the interests protected by privacy rights is complex in the online world. Closed groups and the use of encryption may suggest that private communications are at issue. Yet closed Telegram channels used by terrorists can attract thousands of members, all anonymous strangers to each other; and may be used to share terrorist propaganda that, by its very nature, appears to be incontestably public: Conway, M., 'Online Extremism and Terrorism Research Ethics: Researcher Safety, Informed Consent, and the Need for Tailored Guidelines', Terrorism and Political Violence 2021 Vol.33, No.2, 367-380. If reasonable expectation of privacy is the yardstick, public expectations may vary wildly: Fiesler, C., Proferes, N., '"Participant" Perceptions of Twitter Research Ethics', Social Media + Society (January 2018) found that Twitter users frequently did not appreciate the consequences of public tweeting.
[55] Liberty and Others v The United Kingdom, App No 58243/00, Judgment, European Court of Human Rights (1 July 2008)
[56] Big Brother Watch v UK, App.Nos. 58170/13, 62322/14 and 24960/15 (25 May 2021, ECtHR, Grand Chamber).

consequences. This places inevitably places greater weight on **content removal** as a counter-terrorism strategy.

Counter-terrorism values

1.34. The chief object of counter-terrorism is to safeguard the population from acts of violence which pursue a political, religious, racial or ideological programme[57]. **Physical violence does not take place in cyberspace**[58], so countering terrorism online concerns addressing online content or behaviour that might lead to real world violence on some subsequent occasion.

- Preventing physical violence is the right metric rather than preventing terrorism offending. Not all terrorism offences result in terrorist harm. Rather, they penalise prior conduct, enabling the authorities to intervene before physical violence has taken place. Through speech offences such as encouraging terrorism, or disseminating terrorist publications[59], the law tacitly accepts that using words and images may lead to violence but the words and images are not themselves acts of terrorism.
- Conversely, considering words and images as species of terrorist harm results in a distorted feedback loop, in which restrictions on internet activity may appear justified, even though they make not one bit of difference to violence in the real world. The Online Safety Bill's use of the phrase "illegal content" and "terrorism content"[60] could encourage such thinking.
- If eradicating certain types of content becomes a counter-terrorist goal in itself, then it is difficult to apply an evidence-based approach to whether online restrictions are justified. The same is true if one attempts to redefine violence[61].

1.35. The lack of directness between words and violence and terrorist harm is crucial to understanding why the topic of terrorism online is so difficult.

1.36. *Firstly*, terrorist violence may be **enabled** through plans formulated (acquiring details of targets) discussed (within a terrorist group or cell) or methods obtained (techniques for killing through to 3-D printed weapons[62]) on the internet. There are

[57] Terrorism is defined in section 1 Terrorism Act 2000.
[58] I recognise that in a future metaverse, it might be necessary to give serious consideration to the protection of digital avatars, depending on the importance particularly they play in every day lifeArticle 8(d) of the UN's Internet Governance Forum's 'Charter of human rights and principles for the internet' concerns the right to and inviolability of virtual personalities. Terrorist cyber-attacks are catered for in the UK definition of terrorism (s1(2)(e) Terrorism Act 2000 includes action designed seriously to interfere with or seriously to disrupt an electronic system) but terrorist cyberattacks are some way off: Ciaran Martin, head of National Cyber Security Centre, interview with Wired (6.5.17).
[59] Sections 1, 2 Terrorism Act 2006.
[60] Clause 52(2), (5), Schedule 5.
[61] E.g. DeCook, J., 'Safe from "harm": The governance of violence by platforms' (2022) 14 *Policy and Internet* 63 (referring to "symbolic and cultural violence").
[62] R v Hall, Salmon, Wright and Whibley (Doncaster Crown Court, 2022).

sufficient cases of improvised explosive devices being made to an internet recipe[63] to know that some online material lowers the bar to certain terrorist acts which were previously dependent on expert bomb-makers or the covert circulation of physical manuals. Explosives manuals are not, however, self-executing. An individual must decide to exploit the know-how for terrorist ends.

1.37. *Secondly*, terrorist violence may be persuasively enjoined through a phenomenon that is universally referred to as **radicalisation** but which is barely understood and frequently contentious[64]. Stuart MacDonald and Joe Whittaker have pointed out a serious lack of clarity in the use of terms such as radicalisation, online radicalisation and self-radicalisation[65], and Maura Conway has pointed out that, however one defines it, radicalisation often has a social dynamic, which is at odds with the popular image of passive consumption[66].

1.38. The role played by the expression or consumption of words, images and sounds on the internet, in the subsequent commission of terrorist violence remains elusive, but it is undoubtedly true that most terrorism arrestees are profoundly engaged in expressing and consuming violent and hateful material online. There are patterns of mass ideological violence where the influence of online materials appears incontestable.

- The Buffalo (US) killer Payton Gendron was inspired by Christchurch (NZ) killer Brenton Tarrant who was inspired by the Norwegian terrorist Anders Breivik[67], and so on. It would be irresponsible for the authorities in the UK not to be supremely mindful about similar violence in the UK, especially if effective 3-D printed guns take hold[68].
- Proscribed organisations such as Islamic State would hardly put such store by publicly available content if online materials had no real world consequences: hence as JM Berger has illustrated, their concern when they are driven onto less publicly available channels[69].

---

[63] Gill, G., Corner, E., McKee, A., Hitchen, P., Betley, P., 'What Do Closed Source Data Tell Us About Lone Actor Terrorist Behavior? A Research Note' (2022) 34 *Terrorism and Political Violence* 113 found that evidence of bomb-making manuals was identified in over 70% of their sample.

[64] Faure Walker, R., 'The Emergence of 'Extremism': Exposing the Violent Discourse and Language of 'Radicalisation' (Bloomsbury, 2021) takes issue with the idea that expression of radical beliefs is a predictor of future acts of violence.

[65] Macdonald, S. & Whittaker, J. (2019). *Online Radicalization: Contested Terms and Conceptual Clarity*. John R. Vacca (Ed.), Online Terrorist Propaganda, Recruitment, and Radicalization, Boca Raton: CRC Press.

[66] 'Determining the role of the Internet in Violent Extremism and Terrorism: Six Suggestions for Progressing Research', Studies in Conflict in Terrorism Vol. 40 (2017).

[67] 'Buffalo shooting: How far-right killers are radicalised online', BBC News (17.5.22).

[68] Burgess, S, '3-D printed guns are appearing on British streets – and the police are taking notice' (Sky News, 15.6.22).

[69] Berger, J.M., Perez, H., 'The Islamic State's Diminishing Returns on Twitter: How suspensions are limiting the social networks of English-speaking ISIS supporters', occasional paper, GW Program on Extremism (2016)

1.39. But the truth is that **not only degenerate content** – live-streaming of massacres, pictures of beheadings, violent manifestos – has the capacity to radicalise. Extracts from the Qu'ran are used to exhort terrorist killings[70]. A BBC documentary appears to have malignly influenced a terrorist murderer of a Muslim worshipper[71]. There is no definition in law that can capture material that has the capacity to radicalise, but is otherwise worthy of protection for journalistic, cultural, religious, topical, comedic or other reasons. Nasheeds, even 'jihadi nasheeds', are problematic for this reason[72]. There is particular reason to be sceptical about tech companies deciding what content is worthy of protection, and what is not.

1.40. As far as I am aware there is no estimate of the **eyeballs to violence ratio**: that is, the relationship is between the number of eyeballs on enabling or inspiring content, and the number of terrorist plots or attacks during the same period. It is however possible to be confident that that it is only an exceptionally small subset of consumers who will then go on to use violence[73].

- For a sense of scale, there were 1.5 million video uploads of the Christchurch live-stream in the first 24 hours after the attack[74]. Assuming some degree of automation, and numerous uploads by the same individuals, this suggests a figure of 100s of thousands exposed to this content, which is still available on platforms today.
- An analysis of 33 terrorist-operated websites (including both Islamist and Extreme Right Wing) found 1.54 million monthly visits[75].

1.41. No content is automatically radicalising because the vast majority will respond to terrorist propaganda with aversion. Despite the extension of the UK's counter-terrorism remit to right wing terrorism, and the discovery of huge volumes of hateful and violent online expression, actual terrorist violence remains rare[76].

---

[70] Holbrook, D., Using the Qur'an to Justify Terrorist Violence: Analysing Selective Application of the Qur'an in English-Language Militant Islamist Discourse, Perspectives on Terrorism Vol.4 Issue 3 (2010).
[71] Glazzard, A., Shooting the Messenger: Do Not Blame the Internet for Terrorism, RUSI Newbrief, vol 39 issue 1 (2019); Dodd, V., 'How London mosque attacker became a terrorist in three weeks', Guardian (1.2.18).
[72] Henrik Gråtrud (2016) Islamic State *Nasheeds* As Messaging Tools, Studies in Conflict & Terrorism,39:12, 1050-1070.
[73] For a sense of scale, there were 1.5 million video uploads of the Christchurch live-stream in the first 24 hours after the attack: New Zealand Government, 2021 Digital Violent Extremism Report (at p31). Assuming some degree of automation, and numerous uploads by the same individuals, this suggests a figure of 100s of thousands exposed to this content, which is still available on platforms today (ibid). Berger, J.M., Perez, H., 'The Islamic State's Diminishing Returns on Twitter: How suspensions are limiting the social networks of English-speaking ISIS supporters', occasional paper, GW Program on Extremism (2016) contains an analysis of IS Twitter supporters June to October 2015. Tech Against Terrorism, 'The Threat of Terrorist and Violent-Extremist Operated Websites' (January 2022), found that a sample of 33 out of 198 identified websites had 1.54 million monthly visits.
[74] New Zealand Government, 2021 Digital Violent Extremism Report (at p31)
[75] Tech Against Terrorism, 'The Threat of Terrorist and Violent-Extremist Operated Websites' (January 2022).
[76] Scrivens, R., Examining Online Indicators of Extremism among Violent and Non-Violent Right-Wing Extremists (2022) *Terrorism and Political Violence*: little is empirically known about the differences in the

1.42. Of course legislative action can prove a temptation precisely because of the volumes of material involved: even if only an infinitesimally small group of users will be moved to violence by content, at these volumes that could translate into a material risk to the public. This broad position appears to underlie the Online Safety Bill[77].

1.43. *Thirdly*, online propagation is simply another method by which terrorists can amplify the terrorising effect of their violence. That has always been the objective of terrorism, which is why discourse on terrorism, and the laws we make, must be careful to keep a sense of proportion.

Conclusions

1.44. In the light of this analysis of the applicable values, it is possible draw the following tentative conclusions.

1.45. **Firstly**, wishing **general and permanent** restrictions on the internet freedoms of millions and billions of users, based on the violent actions of the spectacular few[78], seems an unnecessarily heavy price to pay, akin to banning knives or alcohol [79].

1.46. More acceptable may be measures which disrupt patterns of terrorist behaviour. This is sometimes referred to as the process of adding friction by, for example, making it more difficult to re-post at scale[80], or requiring age-verification by children.

1.47. **Secondly**, the **strongest case** for general and permanent removal[81] concerns content that, on the evidence, is most strongly connected to past and future violence, such as the live-streaming of attacks, or propaganda produced by identified terrorist organisations.

1.48. **Thirdly**, in the absence of a workable distinction between content associated with subsequent violent acts that is nonetheless worthy of protection, and content

online patterns of violent terrorists compared to non-violent extremists who share similar ideological beliefs.

[77] "Although it is hard to quantify the benefit of the removal of terrorist content and activity from the online sphere, it's [sic] removal will almost certainly have an effect on the level of terrorism in society": HMG, Impact Assessment (31.1.22)

[78] To use the phrase coined by Mark Hamm for his 2013 book of the same name dealing with terrorism in prisons.

[79] In this vein ECHR cases in which the internet interference was not proportionate under Article 10 include: Ahmet Yildirim v Turkey App.No.3111/10 (18.12.12) (indiscriminate blocking of access to Google); Cengiz and others v Turkey App.Nos.48226/10 and 14027/11 (1.12.15) (indiscriminate blocking of access to YouTube); Kharitonov v Russia App.No.10795/14 (23.6.00) (collateral effects of blocking IP address of shared web-hosting service).

[80] As advocated by the Facebook whistle-blower, Frances Haugen: see for example, oral evidence to Joint Committee on the Online Safety Bill (25 October 2021).

[81] Including bars on uploading.

that is not[82], **decisions by democratic rights-based states** such as the UK are the least bad way of determining which dangerous material should be removed. In the UK there are various legislative tools that could be adapted for this purpose[83].

1.49. **Fourthly**, even where general and permanent removal is otherwise justified, given the uncertainty that content will result in real world violence, there should be **absolute protection for the most responsible media outlets** to deploy any content as determined by editorial judgment. It must be seriously questioned whether the Bill in its current form give sufficient protection to journalistic freedoms and in particular editorial decisions by "recognised news publishers"[84] to tell the truth and inform public debate without restriction.

1.50. **Fifthly**, an **infrangible record** of all content must be kept in the interests of truth and history to which appropriate access must be permitted. This is consistent with information being archived but not made freely available, or accessible only on application by bona fide researchers[85].

Jonathan Hall QC
June 2022

[82] New Zealand's Chief Censor's conclusion on the Buffalo shooting video was "…there is no merit in this": NZ Classification Office, News Item (16.5.22).
[83] Proscription of terrorist groups under section 3 Terrorism Act 2000; takedown notices under section 3 Terrorism Act 2006; and, in some cases, sanctions (cf the Russia (Sanctions) (EU Exit) (Amendment) (No.9) Regulations 2022 that required social media and internet services to take reasonable steps to prevent UK users encountering certain content).
[84] Defined in the Online Safety Bill at clause 50.
[85] As permitted by New Zealand's Films, Videos, and Publications Classifications Act 1993, section 44, in respect of banned materials.