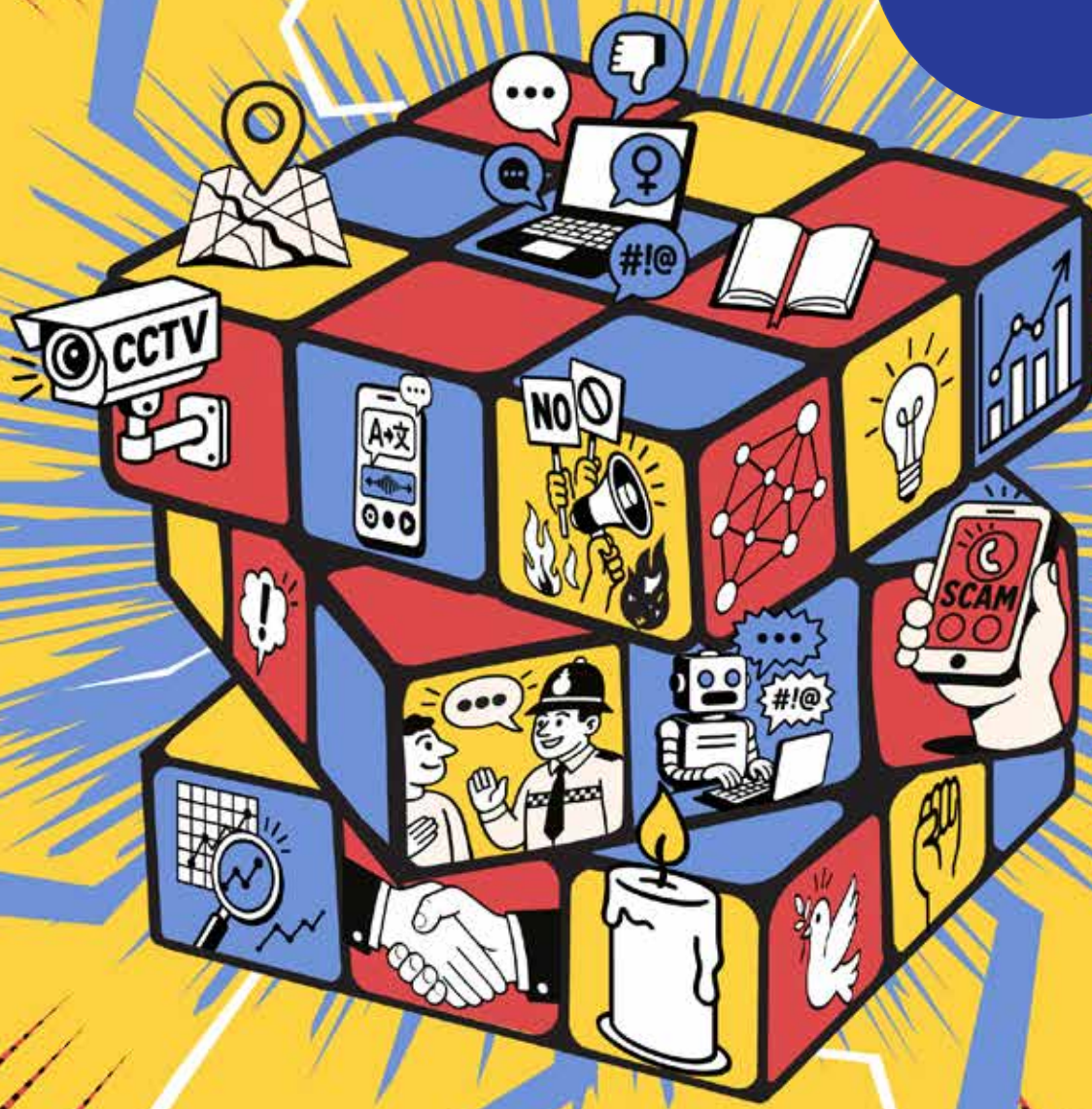
AUTUMN 2025 / ISSUE 21

CREST SECURITY REVIEW

ISSN 2398-0540 (PRINT)
ISSN 2398-0559 (ONLINE)

CSR



System Resilience

HUMAN VS AI TRANSLATION PUTS POLICING RESILIENCE TO THE TEST – p. 12

PLANNED RESILIENCE: WHY STRATEGIES FAIL WHEN THEY FORGET REAL PEOPLE – p. 16

BUILDING COMMUNITY RESILIENCE IN NORTHERN IRELAND FOR LASTING TRANSFORMATION – p. 32

CONTENTS

3 From the Editors

SYSTEM RESILIENCE

4 **Artificial Intelligence and Gendered Security**
AI technologies bring promising opportunities but also new risks, especially for women, girls, and minority communities.

6 **Synthetic Celebs & Fabricated Family: Protecting Against a New Breed of Cybercrime**
With the rapid increase of AI-generated speech, how can the public build resilience against emerging crimes that exploit this technology?

8 **Technology Evolves the Tactics: Preparing for the Rise of Terrorist AI Harms**
How terrorist groups are adopting AI for propaganda, radicalisation, and attack innovation, and how defences are responding.

10 **Transparency Statements Restore Trust & Resiliency**
What should officers say in order to create a collaborative and resilient community?

12 **Interpreting and Translation in the Police Setting: Human vs AI**
Do interpreting and translation services—human or AI—support or undermine resilient policing?

16 **Planned Resilience: Are We Setting Ourselves up to Fail?**
How emergency planning can succeed when it accounts for real people, not idealised scenarios.

20 **Automatically Generating Counter-Speech: Opportunities and Challenges of Using LLMs for CVE**
How are extremist actors exploiting large language models (LLMs) to create and share content.

27 **Why Study Non-Radicalisation?**
How young people in radical milieus resist extremist pathways and build resilience.

30 **Social Networks and Resilience**
What makes a social network less resilient? How structure and behaviour affect vulnerability to threats.

32 **Community Resilience in Northern Ireland: A Framework for Transformation**
How does a community prepare, respond, and recover from disruptive challenges to become more resilient?

36 **The Next Generation of Resilience**
A reflection of the benefits that NextGen networks offer to the study of terrorism and violence.

Beyond the Theme

23 **SPECIAL: CREST Turns Ten!**
A timeline reflecting on CREST's journey over the past decade and celebrating its impact.

38 **Illuminating the Forensic Black Box: Understanding Radicalisation in Secure Settings**
Does our knowledge of radicalisation in the community apply to forensic institutions like prisons or secure psychiatric hospitals?

40 **The Copper's Nose: Spidey Senses Aren't Just for Spider-Man**
Have you ever walked into a situation and instantly known that something just wasn't right?

42 **Read More**
Links and references

CREST SECURITY REVIEW
Editor – Rebecca Stevens
Co-editors – Paul Grasby, & Kathryn Haywood
Layout design – Adele Armistead
Cover illustration – Rebecca Stevens
To contact *CREST Security Review* email **csr@crestresearch.ac.uk**

PAST ISSUES
To download (or read online) this issue, as well as past issues of *CREST Security Review*, scan the QR code or visit our website: **crestresearch.ac.uk/magazine**

FROM THE GUEST EDITOR

System resilience lies at the heart of security. Whether confronting cyber threats, terrorism, or community disruption, the ability for systems to adapt, recover, and maintain core functions is crucial.

In this issue of CREST Security Review, we explore how system resilience is being tested across multiple fronts. From the risks of emerging AI and evolving extremist tactics, to the challenges of building trust in policing and planning for real-world emergencies, resilience is viewed through both technological and human lenses. We also examine the pathways that shape resilience in young people, social networks, and communities.

We begin with Peterscheck (p. 4), who examines how artificial intelligence can be misused in ways that disproportionately harm women, girls, and minority communities. McVean (p. 6) continues the focus on AI with a look at synthetic speech and its role in new forms of cybercrime, while Stevenson (p. 8) explores how extremist groups are already incorporating AI into their tradecraft.

Resilience also depends on how systems build trust and communicate. Dobson (p. 10) shows how a simple transparency statement can help officers foster collaboration and restore confidence in policing. Esposito (p. 12) then highlights the importance of interpreting and translation in legal contexts, raising questions about fairness and accuracy when these services are delivered by humans or AI. Planning, too, is central to resilience: Sonnen (p. 16) warns that strategies written for idealised staff rather than real people risk setting systems up to fail. Rogers (p. 20) turns back to technology, assessing both the promise and pitfalls of using large language models to automatically generate counter-speech against extremist narratives.

Looking at resilience through different pathways, Pilkington (p. 27) argues that studying non-radicalisation helps us understand how young people navigate adversity and resist extremist milieus. Crossley (p. 30) takes this further at the network level, showing how social ties can both strengthen and undermine resilience depending on how they are structured. Hazelden (p. 32) moves to the community scale, proposing a framework for resilience in Northern Ireland that supports not just recovery but transformation. Finally, Squires (p. 36) highlights how NextGen networks foster resilience within the research community itself, creating spaces for collaboration and future preparedness.

Beyond this issue's theme, Henrich (p. 38) reviews the limited evidence on radicalisation in secure settings, revealing critical gaps in our understanding, while Quin (p. 40) examines the "copper's nose," exploring how instinct and professional experience continue to shape decision-making in policing.

Finally, sitting at the heart of this issue, we also mark a major milestone with CREST's tenth anniversary (p. 23), featuring a special timeline that reflects on the centre's journey and impact over the past decade.

As always, we welcome your reflections on this issue. Please share your thoughts through our feedback survey, and let us know what topics you would like to see in future editions. Thank you.

Rebecca Stevens
Editor, *CSR.*

GIVE US YOUR FEEDBACK!

Please fill in the short (and anonymous) questionnaire at this link, or QR code:

www.crestresearch.ac.uk/csr-survey

This questionnaire lists all issues of *CSR* with 3 questions next to each. Please only respond to those issues you have read.

ANNE PETERSCHECK

ARTIFICIAL INTELLIGENCE AND GENDERED SECURITY

Emerging artificial intelligence (AI) technologies bring promising opportunities but also new risks. Among these risks is the potential for AI tools to be (mis)used in ways that disproportionately harm women, girls, and minority communities.

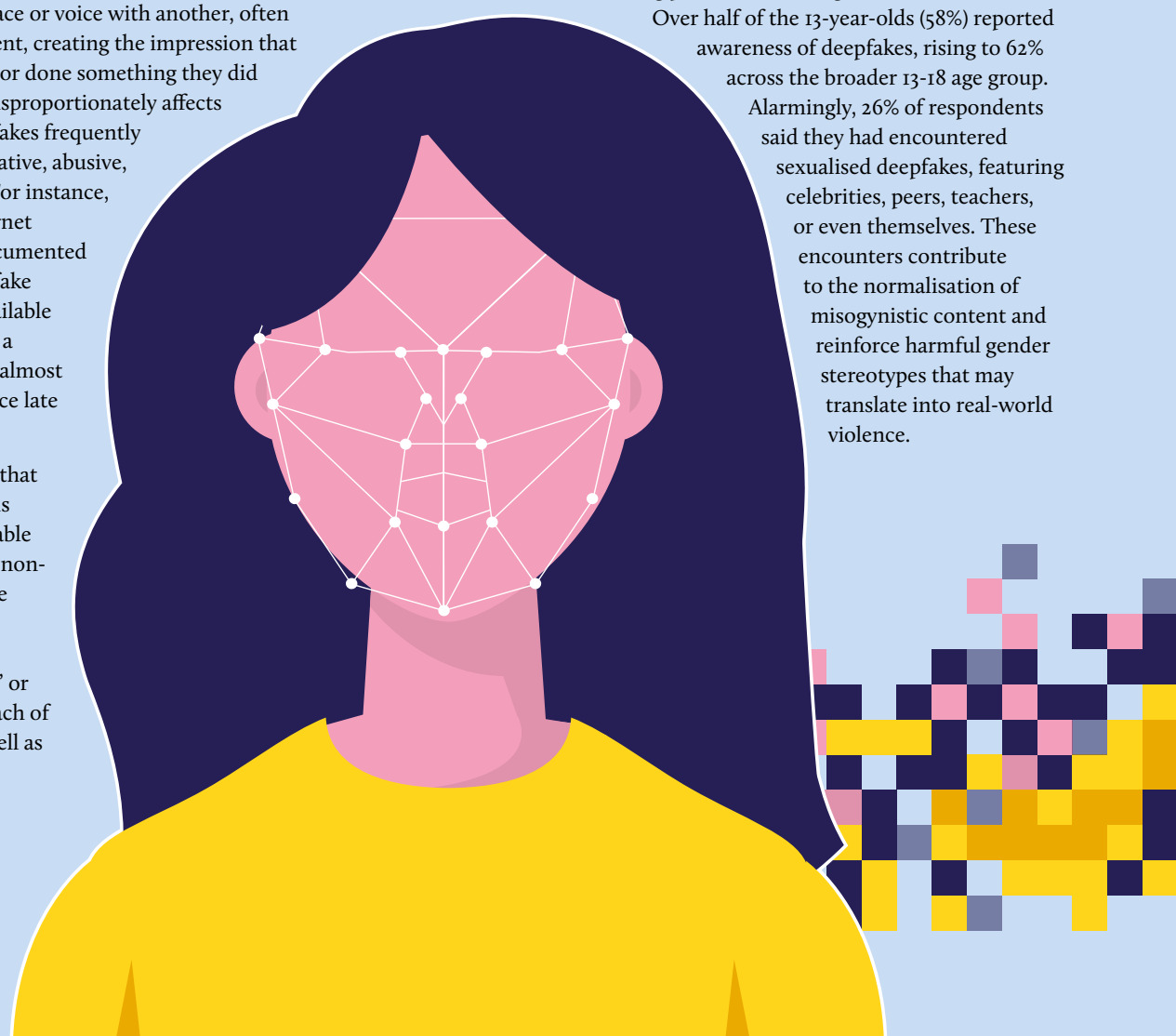
DEEPFAKES

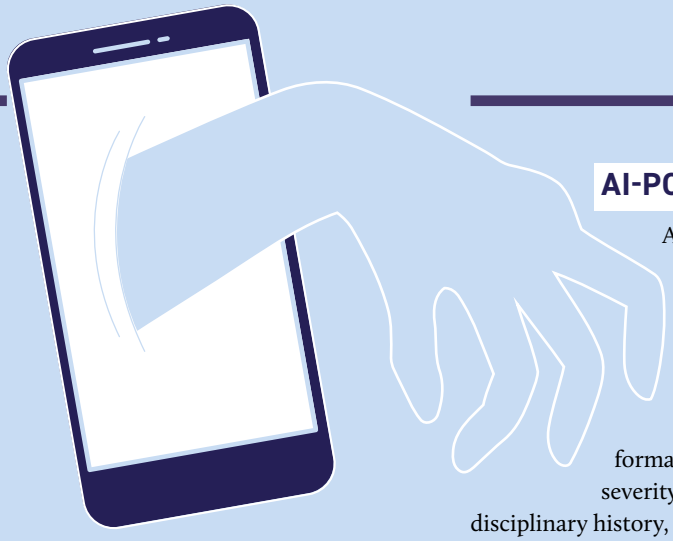
Deepfakes are an increasingly concerning aspect of technology-enabled violence against women and girls. Produced using artificial intelligence, these manipulated images and videos can replace a person's face or voice with another, often without their consent, creating the impression that they may have said or done something they did not. Such misuse disproportionately affects women, with deepfakes frequently deployed in exploitative, abusive, and explicit ways. For instance, a 2025 Oxford Internet Institute report documented nearly 35,000 deepfake models publicly available for download, with a combined usage of almost 15 million times since late 2022.

The authors found that 96% of these models focused on identifiable women to produce non-consensual intimate imagery, and many include search tags such as 'porn,' 'sexy,' or 'nude,' often in breach of platform rules as well as UK legal standards.

YOUTH EXPOSURE AND THE NORMALISATION OF MISOGYNY

According to the Girls' Attitude Survey, deepfakes have become a troubling presence in teenagers' social environment. Over half of the 13-year-olds (58%) reported awareness of deepfakes, rising to 62% across the broader 13-18 age group. Alarmingly, 26% of respondents said they had encountered sexualised deepfakes, featuring celebrities, peers, teachers, or even themselves. These encounters contribute to the normalisation of misogynistic content and reinforce harmful gender stereotypes that may translate into real-world violence.





AI-POWERED HARASSMENT

Artificial intelligence technologies are increasingly implicated in the amplification of online abuse, particularly against women, girls, and racialised individuals, posing serious risks to mental and emotional wellbeing. Recent research shows that individuals from certain racial and gender groups are more frequently targeted, and higher depression, disciplinary issues, and low self-esteem are associated with greater cyberbullying severity. One explainability format ranks race and gender among the most impactful predictors of severity, and their labelling framework explicitly treats bullying, depression, disciplinary history, and heavy internet use as vulnerability factors.

Additionally, companion chatbots, such as Replika, have demonstrated troubling behaviours by, for instance, persistently crossing user boundaries with unsolicited sexual advances, leading to feelings of violation and mistrust among users seeking non-romantic interactions.

To build an inclusive and trustworthy AI ecosystem, consideration of gender must be treated as a foundational principle, not an afterthought.

ECONOMIC INSECURITY

Women face distinct economic risks in the age of AI. A 2024 report by the UNC Kenan-Flagler Business School highlights a pronounced gender disparity in exposure to generative AI automation, referring to tasks that could be performed either partially or fully by AI tools. Specifically, it finds that 8 in 10 working women in the U.S., which amounts to approximately 58.9 million women, are in occupations expected to be highly exposed to AI automation, primarily concentrated in knowledge-based roles. This level of exposure poses significant risks of job displacement.

More recently, the International Labour Organization reports that in high-income countries, 9.6% of female employment falls into the highest risk category for automation, compared to just 3.5% for male employment. Economic security is not just a financial issue but has broader social consequences. For instance, women with reliable incomes are less likely to remain financially dependent on abusive partners, and more likely to exit dangerous relationships. Therefore, automation-induced disruption may deepen systemic risks unless these risks are tackled through intersectional policy interventions.

CONCLUSION

These AI-driven harms reflect existing social inequalities, reinforcing patriarchal and discriminatory logics through automated systems. For victims, and especially those subject to multiple forms of discrimination, the consequences can be deeply traumatic, amplifying systemic exclusion across both digital and physical spaces. Gendered economic insecurity compounds this vulnerability as automation disrupts female-dominated sectors causing a rise in precarious employment, meaning many women face reduced financial autonomy and limited access to recourse.

To build an inclusive and trustworthy AI ecosystem, consideration of gender must be treated as a foundational principle, not an afterthought. Without this, AI risks deepening existing inequalities and falling short of its potential to support a more equitable and democratic society.

Dr Anne Peterscheck is a researcher at the University of St Andrews.

HOPE MCVEAN

SYNTHETIC CELEBS & FABRICATED FAMILY: PROTECTING AGAINST A NEW BREED OF CYBERCRIME

AI-generated speech is a synthetic data type posing a serious and current risk, increasingly used in crime. This article introduces ongoing efforts in increasing public resilience against AI-mediated crime.

Recent technological developments mean that AI-generated speech has skyrocketed in quality, to the point where it is in some cases virtually indistinguishable from genuine human speech. Even more concerningly, speech can be specially engineered to sound like a specific speaker, a phenomenon known as voice spoofing. Spoofs (also known as voice clones or auditory deepfakes) can fool automatic speaker verification (ASV) systems used in secure settings such as banking and government, potentially granting fraudsters access.

Human listeners are similarly at risk. To best illustrate this, we turn now to the 'the 3 Ds of Deception':

DISINFORM – in which spoofs are used to instigate and disseminate knowingly false information. In successful cases, listeners believe and further spread the information, because to all appearances, the source is an influential, ostensibly trustworthy figure. This D is most commonly seen in politics, where cybercriminals impersonate people in power to suppress voters or manipulate elections. At their worst, spoofs of this nature have the potential to threaten democracy itself.

DEFAME – unlike the other two Ds, the main victims here are those being spoofed, not the listeners being deceived. Defamatory voice spoofs create fake audio of their victim saying something offensive, inflammatory, or illegal, triggering social and sometimes financial backlash. While high-profile cases exist, the risk to ordinary people can be greater. A celebrity might weather the backlash; a working-class shop assistant might not. A deepfake of the latter making a racist remark would not spark the same levels of international outrage as, say, one of Billie Eilish would. However, if both lost their jobs as a result, one would lose their only income, while the other would still be worth $50 million.

DEFRAUD – voice spoofs can be used to impersonate someone with a relationship to a listener, with the aim of extracting money from that listener. AI-mediated fraud is a new type of confidence scam, exploiting pre-existing relationships and trust, rather than building trust over time. Fraud can involve false celebrity endorsements or phone scams, in which an imposter poses as a colleague, friend, or family member in financial need. This D has already shown to be hugely damaging, with millions of pounds being stolen this way in recent years (see Stupp, 2019; Brewster, 2023; Bunn, 2023; Verma, 2023).

Presently, attempting to protect human listeners from AI-mediated deception is difficult, since the conditions dictating accurate discrimination between genuine and synthetic speech are little understood. There is, however, one factor that may prove invaluable to increasing this understanding: familiarity. As demonstrated above by all 3 Ds, spoofed voices are celebrities, colleagues, family members – people who are known to the listener. Listeners will have a developed knowledge base of the spoofed speaker's genuine voice. Surely, then, this familiarity will make listeners better at recognising when the speaker is being impersonated?

The conditions dictating accurate discrimination between genuine and synthetic speech are little understood.

Our preliminary findings investigating the influence of familiarity on detection of deepfake voices suggest that this is not the case. Familiarity may degrade performance. Of 120 participants studied, it was found that accurate identification of AI-generated samples dropped to 63% if the perceived speaker was a very well-known celebrity (Donald Trump), compared to one with a more limited fanbase (Linus Sebastian), where accuracy rates were around 73%.

One possible explanation for this is that familiarity with someone's voice isn't limited to just knowing how they sound. The voice activates a wealth of rich information about that person: knowledge of their experiences, personalities, values, and relationships. In the study above, many listeners left comments indicating strong negative feelings for Donald Trump. One listener described the speaker they were listening to as an 'idiot', and therefore decided it must genuinely be Trump, rather than a spoof. In short, they used a highly subjective value judgement to discriminate between human and AI-generated speech, rather than the voice alone.

This suggests that a listener's personal experience with the speaker might overshadow objectivity in decision-making. Compare this to a stranger's voice – the listener has had no past contact with them, and so they can assess the voice impartially.

Future work in this research series will diverge from familiar-celebrity spoofs in favour of spoof detection when there are familiar-intimate relationships between listener & perceived speaker (i.e., family, friends). From the insights gained – particularly why listeners make their decisions – we can explore the cues that are being registered, and more crucially those that are being missed. Spoofing is not an isolated threat; people and organisations at all levels of society are potential victims. Understanding how and why people are deceived is essential if we wish to develop our resilience against them.

Hope McVean is a current PhD student at Lancaster University, working primarily in FACTOR (Forensic Linguistics, Cybersecurity and Technology Research).

JAMES STEVENSON

TECHNOLOGY EVOLVES THE TACTICS: PREPARING FOR THE RISE OF TERRORIST AI HARMS

Terrorist groups, like the societies they emerge from, adapt to new technologies. As AI capabilities evolve, so too do the tactics of extremist actors. While the full effects may take years to observe, as the technologies continue to develop, we are starting to see them directly alter extremism tradecraft.

From AI-driven denial-of-service attacks and adaptive malware, to drone swarms, and autonomous vehicles - the malicious use and wider adoption of machine learning and artificial intelligence, by terrorist groups, has been widely speculated in recently months. Earlier this year, the UK Government published "The Terrorism Acts in 2023 report of the Independent Reviewer of Terrorism Legislation" a report, which among other areas, considers seven categories of potential terrorism harm that may result from use of Generative AI.

This article builds on these conclusions, providing real world examples of terrorist uses of AI.

PROPAGANDA AND PRODUCTIVITY INNOVATION

Synthetic propaganda is nothing new, deepfakes (audio and visual) have historically been used to support terrorism. With the increasing availability of generative AI tools, we're seeing its weaponisation by terrorist organisations for propaganda generation. ISIS' media division, News Harvest, have adopted generative AI to produce human-like news presenters resembling shows such as CNN and Al Jazeera - generating video, audio, and text content tailored for propaganda in various languages.

CHATBOT RADICALISATION

With the boom of generative AI, people have been flocking to these AI chatbots to converse and share ideas. These chatbots provide a real potential threat of steering users towards terrorist ideals and views - whether in closed-loop private one-to-one chats or in public multi-user settings.

We've recently seen this phenomenon with the xAI chatbot, Grok. On 8 July, 2025, Grok began posting content praising Adolf Hitler, using antisemitic stereotypes, and even referred to itself as "MechaHitler". Similarly, before these concerns on X, Gab's AI chatbot had been observed generating Holocaust denial and other conspiracy theorist content.

ATTACK FACILITATION AND INNOVATION

Generative AI may be used to obtain, or refine practical instructions, training, or support for acts of terror. Such tactics are increasingly common in pro-IS propaganda networks. For example in March 2025, a pro-IS account released an AI-generated video featuring a digital avatar providing instructions for making a bomb using common household items.

Generative AI can also be seen to provide innovative recommendations for new types of attacks. On 6 December 2022, a frequent user of a Rocket.Chat server run by Islamic State published a post claiming to have used the free version of ChatGPT to ask how best to "establish the Caliphate." According to the user, the AI responded with a series of operational steps, which he claimed showed it was "smarter than most activists." He also asserted that the response was original and included the full text of the alleged reply.

MODERATION EVASION

Recently the International Center For Counter Terrorism released findings observed from far right clusters on mainstream chat-based platforms (Reddit, Telegram, X, and 4chan) where users shared tactics for generating harmful yet lawful AI content purposefully designed to not be flagged by content moderation systems. Evasion methods included: platform switching (users temporarily migrated from mainstream to fringe platforms to create harmful content, then returned to mainstream sites to spread it more widely), self-censorship (users deliberately avoid sharing technical detail to reduce the risk of content removal), delegation (users ask others to perform tasks that could get them flagged), banishment sharing (users discuss their past bans to inform others about platform moderation limits), censorship evasion education (banned users share strategies

Chatbots provide a real potential threat of steering users towards terrorist ideals and views - whether in closed-loop private one-to-one chats or in public multi-user settings.

to continue their activities elsewhere, effectively teaching others how to bypass restrictions), and finally tool and site recommendations (individuals suggest alternative platforms or tools with fewer safety or moderation measures).

SOCIAL DEGRADATION

Finally, this category covers AI driven content that spreads distrust, conspiracy narratives or social division without directly urging terrorism.

We see this behaviour in the events less than three hours after the Southport murders. Where on 29 July 2024, the X account Europe Invasion shared an AI-generated image showing knife-wielding men in traditional Muslim dress outside Parliament, next to a crying child in a Union Jack T-shirt. The post amassed around 900,000 views and introduced the slogan “Protect our children”. Similarly, an anti-immigration Facebook group created an AI-made image inviting people to a rally in Middlesbrough depicting a large crowd at the cenotaph. Other tools, like Suno, were used to produce xenophobic songs such as “Southport Saga,” featuring AI-generated vocals stating things like “hunt them down somehow”.

Meanwhile Tech Against Terrorism traced a brand-new TikTok account that began posting only after the Southport attack. Despite having no prior history, its AI-generated protest posters quickly reached 57,000 views, a spread the group links to coordinated bot amplification networks.

CONCLUSION

The adoption of AI by terrorist groups is here, it’s not going away, and will only continue to amplify terrorist tradecraft over time. However, our defensive toolkit is evolving also: platforms are now beginning to auto-watermark AI-generated translations and graphics (such as Meta’s “Imagined with AI” labels and embedded metadata). AI red-team simulations stress-test counter-terror guardrails against bomb-making or cyber-threat scenarios; shared-hash moderation coalitions, powered by GIFCT and Hasher Matcher Actioner (HMA), now synchronise detection of mutated extremist media across platforms; and in education, pre-bunking curricula like the “Bad News” game paired with browser AI fact-checkers have reported cuts of susceptibility to disinformation. These defences, are already being piloted, tested, and put to practice - and they’re only the beginning.

James Stevenson is a security researcher with a decade of experience in the computer security and research industry. He is currently a PhD candidate at the University of Bristol working at the intersection of machine learning, computer science, and social science. His research focuses on detecting and predicting extremist content online, and understanding how extremist groups use AI to enact harms.

KYLE S. H. DOBSON

TRANSPARENCY STATEMENTS RESTORE TRUST & RESILIENCY

Trust is challenging to build but essential for creating a collaborative and resilient community against public safety threats. Kyle Dobson explains how an evidence-based solution begins with an authentic transparency statement.

> **Officers should communicate an honest reason for the interaction that is ideally motivated by helping the community and specific individuals.**

THE ILLUSION OF TRUST

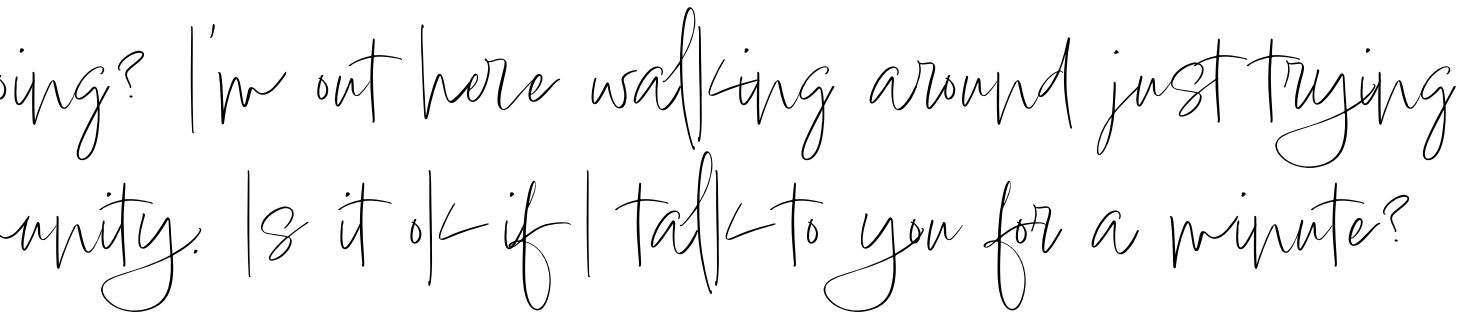
If you're an officer reading this, you might think about how you're actually quite good at building trust with the community. If that's true, you're like the community-oriented police I learned from. However, science calls for accuracy. Do you or others ask the community how much they trust you to help their community stay resilient to threats after your interactions?

I followed up with police and community members after their interactions, independently, to get a sense of how they felt towards each other. Research unveiled an illusion of trust: police thought they earned more trust with community members than they actually had. Officers feel they're being respectful, while community members interpret their interactions as threats. This makes sense. If your boss stops in briefly to check in on you, you probably smile, nod, maintain eye contact, and speak in a friendly tone that sounds like someone who trusts their superior.

TRANSPARENCY STATEMENTS REBUILD TRUST

Our lab's new research shows that a simple intervention to transparently connect with the community can quickly put people at ease: a short statement from the officer about why they stopped the person and what they intend to do. Police can use the first moments to mitigate threat, prevent escalation, and build trust.

Police start with a trust deficit that can feel demoralising for officers who want to build rapport. Strategies that create comfort in anxious situations – like asking permission – can backfire with power imbalance. "You mind if I talk to you for a minute?" sounds scary with unclear intentions.



THE SOLUTION

Give a transparency statement. It's a sentence an officer gives to start interactions with community members – from traffic stops to noise complaints. The statement quickly explains intent. It seems simple, but our work shows a transparency statement can change an interaction.

A transparency statement is not an exact script. Officers can and should word their transparency statements in a way that captures their true goal for the interaction.

For example, "Hi, I'm Officer [Name], how's it going? I'm out here walking around just trying to get to know my beat and my community. Is it OK if I talk to you for a minute?" In this opening statement, the officer states that their intent is to get to know the community, not to hurt them.

Four Elements of Successful Transparency Statements

There are four key elements that we've pinpointed for effective transparency statements.

1. **The first is timing.** The statement should be made as soon as possible, to set the tone for the interaction from the outset.
2. **Second is benevolence.** Officers should communicate an honest reason for the interaction that is ideally motivated by helping the community and specific individuals.
3. **This only works if it is sincere**. Reputation can be ruined if you are inauthentic.
4. **Last, the statement needs to be personal.** Officers should speak in the first-person (e.g., "I'm worried about your safety") and refer to the current situation. Generic statements about department-wide efforts might work less well (e.g., "Our department has a new initiative to get to know community residents").

DATA

In our field studies, transparency statements have a simple but powerful effect. In one experiment, we measured electrical signals given off by participants' skin, which indicate stress, during interactions with police officers. We also analysed the language spoken during the exchanges and surveyed participants after the interactions.

When officers implemented transparency statements, community members were more likely to respond using language associated with positive rapport and trust-building.

When officers implemented transparency statements, community members were more likely to respond using language associated with positive rapport and trust-building. They were more engaged and spoke more during the conversation. Without a transparency statement, skin results showed stress levels rose across the conversation.

In surveys after, community members were less threatened: 40 percent with transparency statements felt low-threat versus 29 percent without. They reported higher trust after transparency statements, too, expecting the police were more invested in their well-being (55 percent versus 46 percent).

Many adults in the United States encounter the police each year. For 20 million Americans, the interaction is initiated by a police officer, typically through a traffic stop or street contact. If officers just across the USA consistently used transparency statements, 1.8 million more people could have trusting interactions with police and 2.2 million more people could have an interaction where they don't feel threatened. One more finding that surprised even us: above most other emotions, transparency statements even made community members feel more inspired after their interactions.

Teaching the method takes minutes, training it takes a few hours, and practice makes it habitual.

Transparency statements are a simple concept, and that's part of their beauty. Law enforcement deals with a high cognitive load in the field among uncertainty. Protecting communities requires collaboration to stay resilient. Having a simple template for transparency can help create the trust required to facilitate police-community collaboration towards a resilient community.

Kyle S. H. Dobson is an assistant professor of public policy and psychology at the Frank Batten School of Leadership and Public Policy at the University of Virginia

GRETA ESPOSITO

INTERPRETING AND TRANSLATION IN THE POLICE SETTING: HUMAN VS AI

Resilient policing systems can adapt, maintain core functions, and uphold fairness even in complex, high-pressure situations. But one often-overlooked factor in that resilience is the ability to communicate accurately across languages.

From major incidents involving multiple nationalities to everyday interviews with non-native speakers, the quality of these services directly affects evidence integrity, decision-making, and public trust. Whether human or AI, poor quality interpreting or translation outputs in high-stakes scenarios can introduce vulnerabilities that undermine investigations, delay justice, and erode confidence in the system.

Imagine finding yourself in a foreign country, and having to deal – at some point – with the justice system of that country, in a language you neither speak nor understand, and the only way to communicate and be informed is via an interpreter or translated documents.

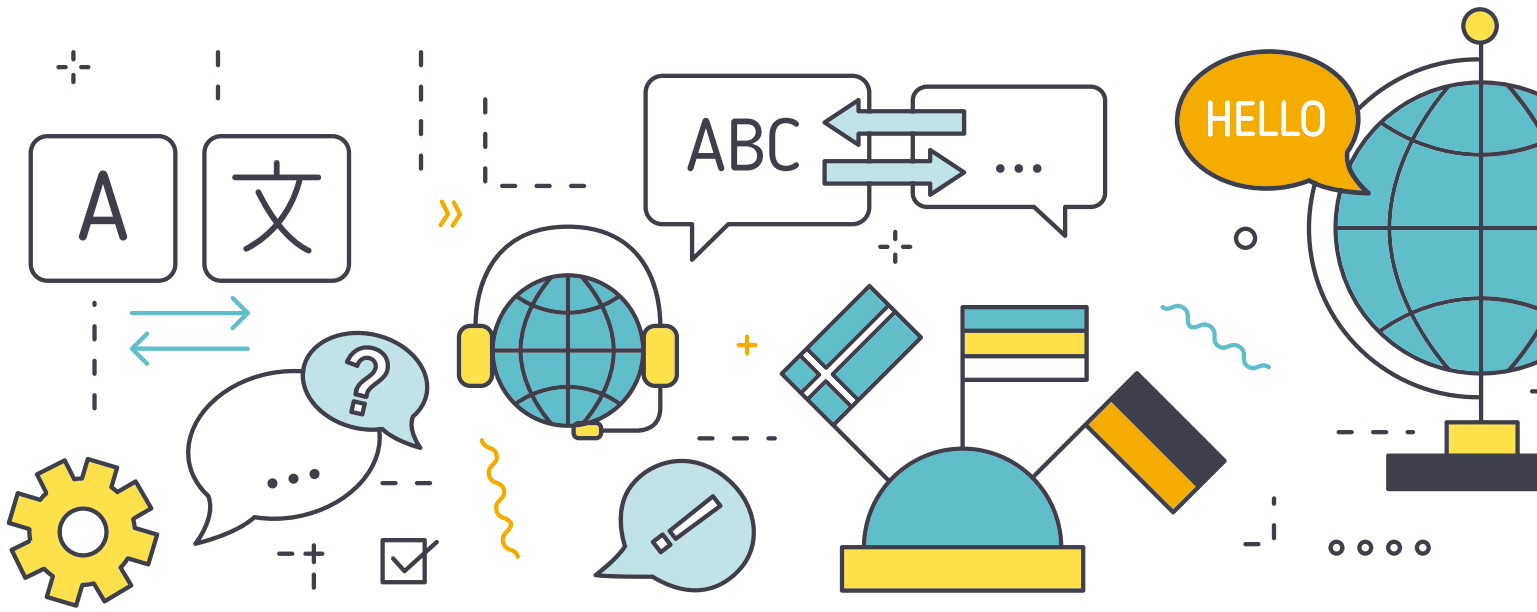
This can be considered a leap of faith, as there is no other way to have direct access to that particular system. Therefore, the provision of high-quality interpreting and translation services, both human and/or AI, needs to be prioritised to guarantee fair justice.

Focusing in on interpreting services in the legal setting, previous research has already established the active role of interpreters in interpreter-assisted interactions, where they can change the course of the legal proceedings.

Nevertheless, especially in the police setting, mainly due to difficulty of access, it is still not clear to which extent interpreters affect conversational dynamics and mingle with interviewing strategies, the quantity and quality of evidence retrievable, and the delivery of justice as a whole.

Furthermore, the spread of AI across the board has brought about new questions to address and new challenges to face. Regarding multilingual interactions, can AI software be a potential way forward and meet the high standards that are vital in policing?

My PhD research at Lancaster University aims to address these issues by means of linguistic analysis, and specifically taking a Pragmatics and Conversation Analysis (CA) approach.



Preliminary findings from the analysis of two interpreter-assisted police interviews in the language pair English-Italian have shown that interpreters do have an impact on a variety of dimensions, including rapport building, interviewing strategies, the perceived credibility and reliability of the interviewee, and the quantity and quality of the evidence obtained, and additionally, that AI software outputs need human post-editing.

In regard to the accuracy analysis of human interpretation, a data-driven, cross-cultural pragmatic approach was taken, where the intended meaning of the input utterance of a primary speaker in the source language is compared against the intended meaning of the output utterance in the target language as rendered by the interpreter. The focus is on the discrepancies between input and output utterances, and recurring modifications in some pragmatic markers have been identified and classified, such as hedges and enhancers, mood, tense, reported speech, presuppositions, and many others.

For instance, if the interviewer's original question is "Oh, e di Amanda che ha detto?" (So, what did she say about Amanda?), and the interpreter's rendition is "Did she say anything about Amanda?", the presupposition behind the original question undergoes a shift, as the original presupposes that the person speaking indeed said something about Amanda – so the question probes about the specific content of this interaction – while the interpreter doesn't presuppose this, they just ask whether the person speaking said anything about Amanda. This modifies the interviewing strategy adopted by the interviewer, which is originally more leading.

WHAT ARE THE KEY DIFFERENCES BETWEEN MONOLINGUAL AND BILINGUAL INTERACTIONS IN THE POLICE SETTING?

The role of interpreters in interpreter-assisted interactions has been studied extensively in different settings, such as legal, medical care, education, asylum seeking, and many others.

Specifically, in regard to interpreter-assisted police interviews, why is this type of interaction different from a monolingual one? What should we pay special attention to?

Firstly, contrary to the neutrality myth surrounding the role of interpreters, they are active participants in the interaction, influencing conversational dynamics in various ways.

For instance, a primary speaker may directly address the interpreter instead of the other primary speaker, or the interpreter might engage in side-conversations with them, ask clarification questions, interrupt, or clarify points. This is because interpreting entails much more than just relaying the content from one language into another.

Pragmatic Markers

Pragmatic markers refer to every element of language that carries pragmatic force. Pragmatic force encompasses the speaker's intended meaning and the effect of an utterance on the listener, which goes beyond the literal meaning.

Interpreting entails much more than just relaying the content from one language into another.

This suggests that interpretation always involves intervention, and we should be focusing on unnecessary types of intervention, but especially being aware of the inevitable shifts that may occur via the interpretation process, without condemning them.

Secondly, the increased vulnerability of primary speakers should be taken into account.

If we consider the canonical arrangement of interpreter-assisted police interviews, it typically involves interviewers who are native speakers of the dominant language in a given country, an interviewee who is a non-native speaker of that language, and an interpreter, the bilingual participant.

In a police setting, the conversational power among interviewers and interviewees is usually unbalanced, due to the nature of the interaction, and this imbalance is exacerbated when the interviewee does not speak or understand the dominant language of the interaction.

Therefore, this augmented vulnerability should definitely be considered when engaging within interpreter-assisted interactions and should result in tailored interviewing strategies.

For instance, it is well-documented that negative polar questions (yes/no questions) pose many issues in multilingual interactions. This is because people from different linguistic and cultural backgrounds may reply with the same answer (affirmative or negative), but intend opposite meanings. For example, with the question "Didn't they come to the party?", in English, "Yes" means "They came", while "No" means "They didn't come". In Italian, you might use "Yes" and "No" to say that the person came, maybe specifying "Yes, they came", or "No, they came".

It is clear how in an interviewing setting, this may be problematic. Therefore, acknowledging these cross-linguistic discrepancies can help refine interviewing strategies, for instance, by avoiding negative polar questions and instead phrasing them in a more neutral form.

Lastly, unique challenges are faced by legal and language professionals when collaborating, due to their different priorities.

This is why proper training should be ensured to facilitate seamless collaboration, with open communication as one of its key pillars. For instance, briefing and debriefing practices pre and post interview can be considered an example of good practice.

In light of this, priority should be given to properly supporting the collaboration between legal and language professionals so they can deliver high-quality service, and fostering academic research to bridge the gap between academics and practitioners, working towards the shared goal of enhancing current policing practices.

> **Priority should be given to properly supporting the collaboration between legal and language professionals.**

Preliminary findings suggest that human post-editing is still required when it comes to the use of AI translation software.

IS AI A VIABLE ALTERNATIVE IN HIGH-STAKES SCENARIOS?

Preliminary findings from my research study suggest that human post-editing is still required when it comes to the use of AI translation software.

For instance, to test Google Translate (GT) – which is a Neural Machine Translation (NMT) software, specifically trained on bilingual data – I fed it with the transcripts of the two interpreter-assisted police interviews to draw a comparative analysis between the primary speakers' inputs, the GT outputs, and the interpreters' renditions.

It emerged that in discrepant GT outputs, severe alterations could be identified that could lead to serious consequences for individuals.

In one case, the original input utterance reads "No, like, it could have been me, like the murder could have happened to me." And GT translated it into Italian as "No, tipo, potrei essere stato io, tipo l'omicidio potrebbe essere successo a me." ("No, like, I could have done it, like the murder could have happened to me.") It is evident how "it could have been me" and "I could have done it" have different responsibility implications in a police interview setting.

Similar conclusions have been drawn in regard to AI speech-to-text software in legal settings, where its use was recommended to draw initial first drafts only.

Nevertheless, my future research aims to understand how AI translation software can be optimally applied to high-stakes scenarios to speed up the proceedings, but without prejudicing the quality of the outputs.

For policing and justice systems to be truly resilient, they must plan for sustained, high-quality multilingual capability.

This means investing in skilled human interpreters, integrating AI tools cautiously into linguistic and policing practices, and ensuring robust quality controls for both.

Building resilience also means creating redundancy so when one channel is unavailable or under strain, another can step in without compromising accuracy or fairness. As global mobility, migration, and technology continue to reshape the policing landscape, interpreter-assisted communication should be recognised as a core operational strength that can speed up proceedings, without prejudicing the quality of the outputs.

Greta Esposito is a PhD student at Lancaster University in the Department of Linguistics and English Language. Her research interests focus on police interpreting, legal translation, and applications of AI software in legal settings. Her Twitter (X) handle is @GretaEspos1997

KT SONNEN

PLANNED RESILIENCE: ARE WE SETTING OURSELVES UP TO FAIL?

Plans written for delivery by an idealised concept of the person completing the job role, rather than a living breathing being, may be overlooking fundamental flaws.

As the Roman philosopher Seneca the Younger declared "errare humanum est", meaning that in spite of our best efforts people are fallible beings. Whilst action plans vary in complexity and efficacy, very few seem to make allowances for the human factor in their delivery, which may lead to their operational failure.

'Plans' in this context refers to a set of protocols or actions to be taken, whether to complete business as usual or in emergency situations. They are found in every business and industry, ranging from simple fire evacuation plans to multi-agency emergency procedures. Whilst this framework of consideration (described below) originated from observing and working within the events industry, its application transcends any individual sector. The aim is not to cast aspersion, but to highlight the differences between the idealised world that the plans are generally written for (populated entirely by highly motivated, well-trained, perfect staff) and reality.

Plans themselves are typically written by those towards the top of the organisational pyramid with a specific field of oversight and knowledge. They are reliant on effective delivery at every level, usually by an idealised concept of the person completing the task, and may include roles the writer may not have personal experience of, or which have changed since they performed them. This form of affinity bias may mean that abilities could be overestimated, again risking the successful completion of the plan and therefore the resilience, both in terms of initial response and subsequent recovery.

What this author proposes is a human approach to identify the person points of potential failure in a plan via mentally examining the foreseeable factors and their impacts through a framework of questions. This should be done for each 'level' of the plan delivery, both 'standard' and 'emergency', to anticipate and address issues that may arise at any job role. Whilst it may not always be the direct responsibility of the plan writer to fix, anticipating possible vulnerabilities allows for failsafes or alternative measures to be built in. These five areas for consideration, reminiscent of Maslow's Hierarchy of Needs, allow for a holistic appraisal of unintentional threats to success, including some potential vulnerabilities to insider risk. Each of the five areas of consideration (Knowledge, Practical, Physical, Emotional, and Motivational) can be examined through the lens of a question, which are individually explored below. Although initially inspired by the events industry, where the impacts may be starkly visible especially set against the backdrop of expectation, this mental examination can just as easily be applied to plans for any industry or sector. Even within a sector, different job roles will likely yield different influences, even if operating on the same site. For example, a control room operator is likely to have the majority of their physical and practical needs met, unlike an individual working outside on their feet in all weathers, however may be more emotionally impacted by the nature of the work itself.

Do you have a plan?
NO
Write one and start again
YES
Knowledge: Does the individual know and understand their role and responsibility as detailed within the plan?
NO
YES
Make appropriate changes and re-evaluate.
Practical: Are they able to fulfil the tasks as detailed within the plan?
YES
NO
Physical: Have their basic functional needs been met?
YES
NO
Sub-optimal delivery of the plan. It may not succeed.
Emotional: Do their working conditions meet their emotional needs?
NO
YES
Motivational: Are they motivated to fulfil their role and responsibility as detailed in the plan?
YES
NO
Your plan may succeed!

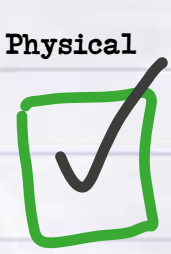
DOES THE INDIVIDUAL KNOW AND UNDERSTAND THEIR ROLE AND RESPONSIBILITY AS DETAILED WITHIN THE PLAN?

It is imperative that instructions are detailed in a clear manner, free from excessive details, complicated jargon, and allowing for differences in language and understanding. Appropriate time should be allocated to get to grips with the briefing, and confirmation of understanding evidenced. Whilst printed handouts may satisfy the minimum legal requirement, without facilitating time for absorption they are unlikely to be more than literal bin fodder, risking potentially sensitive information.

ARE THEY ABLE TO FULFIL THE TASKS AS DETAILED IN THE PLAN?

If the plan requires equipment to be used, this must be provided, accessible, fit for purpose, and maintained, alongside the provision of appropriate training. The assigned individual must also be physically (and mentally) capable of completing the task, requiring consideration of temporary and permanent impairments, and reasonable adjustments, whilst also balancing legal rights.

On a more granular level, there may be additional qualities that are required in particular roles; for example, a meek, indecisive individual may not be suited to assume operational command of an incident. The role itself must be realistic and not ask too much of an individual, especially with regards to risking their personal safety.



HAVE THEIR FUNCTIONAL NEEDS BEEN MET?

Despite widespread evidence of the negative impact of poor-quality, or overall lack of, sleep on both physical and mental capabilities, many industries still rely on overly long shifts for multiple consecutive days or weeks. It is imperative to consider the differences between 'paid time' and the actual time demands of the role (including overtime, travel, signing in/out), as well as the quality of sleep available.

Transparency regarding food quality and availability may allow individuals to make their own arrangements if required. A packed lunch bag may sustain a child for a school day, but not an adult doing 12+ active hours. Is hydration accessible for those on static positions? And toilet provisions?

With weather conditions becoming more extreme, provision should be made for the physical impacts of (prolonged) exposure. Finally, some attention should be paid to the task itself, and keeping the mind engaged and ready to (re)act. These may also be key impact factors in staff attrition, increasing recruitment and training costs.

Emotional

DO THEIR WORKING CONDITIONS MEET THEIR EMOTIONAL NEEDS?

Ensuring staff feel valued, trusted, and empowered may reap rewards beyond what could be attained with financial motivation. Being recognised as an individual is key, part of which is feeling that one's beliefs and values are being accommodated. In times of hardship, this nod towards cohesion through shared social identity may encourage staff to go above and beyond what is expected.

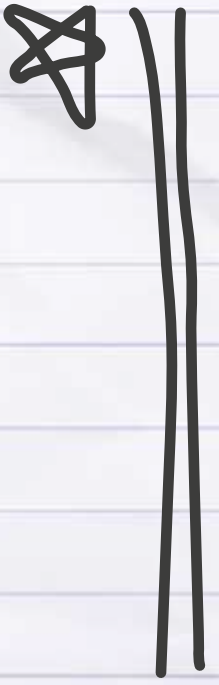
When considering resilience, certain sections of society are already displaying higher levels of resilience, at least outwardly, due to additional pressures around their identity such as skin colour, sexuality, and other aspects of intersectionality. These individuals may be at 'resilience capacity' – is it fair to ask them to shoulder additional emotional burdens? Disparity between individuals, each carrying different psychological loads, means that events and after-effects will impact us all differently. In reality, there is no quick or practical way for this to be assessed, so for the purposes of this framework this is more for additional consideration and understanding. Awareness of individual resilience and pressures within a team may influence deployment choices.

Motivational

ARE THEY MOTIVATED TO FULFIL THEIR ROLE AND RESPONSIBILITY AS DETAILED IN THE PLAN?

Although history shows us that people, in times of need, will generally co-operate and assist others, assumed public goodwill should not be relied upon within resilience planning, especially in lieu of a capable and motivated workforce. Motivating staff can be challenging, with no singular magic formula, however enabling an actuated sense of purpose, ensuring they understand their importance in the hierarchy of action, fostering mutual respect, and supporting ambition will help. Contractors providing well-treated, able, and motivated staff who will step up in the event of a crisis will likely be expensive, however someone unwilling to complete their role may induce a far greater cost.

A good and positive company culture will foster a higher level of motivation. For example, the NPSA authored a useful interactive guide for motivating staff. Although it is designed for those responsible for security officers protecting the critical national infrastructure, the information and suggestions are widely applicable.



Whilst the testing and exercising of plans and measures are typically costly endeavours, mentally evaluating these procedures and the influences upon them from the point of view of the typical staff member in that role may allow for identification of points of failure and subsequent contingencies in an environment where failure is safe and cost-free. Of course, this is reliant on the plans themselves being appropriately detailed, removing assumption and ambiguity, and complementary in the case of different organisations working together.

The mental evaluation of the roles and the individuals completing them should be done by those directly responsible for them, typically a manager, but it would also be an advantageous exercise for those with overall responsibility. It should be highlighted that whilst this could be used at an individual level, the aim is more for general awareness of the challenges that may be faced and how they could impact the safe and effective delivery of the plan.

From a legal perspective, the concepts of foreseeability and reasonably practicable become especially key when a plan fails. Especially given learnings from prior investigations and inquiries, the questions shifts to whether organisations can afford not to explore the ideas raised here. If the safe evacuation of thousands relies on the alertness and enthusiasm of a minimum-wage untrained steward who has been dropped at a remote exit with no contact/food/shelter/water and minimal sleep for several days, it may be time to evaluate not only our expectations but also the system as a whole.

KT Sonnen is a PhD researcher at the Centre for Peace and Security, Coventry University, looking at the safety and security of music events.

ELLIE ROGERS

AUTOMATICALLY GENERATING COUNTER-SPEECH: OPPORTUNITIES AND CHALLENGES OF USING LLMS FOR CVE

As technology has developed, extremist actors have found new ways to use it to their advantage. Large language models (LLMs) are one area that has been exploited by extremists to create and share content.

INTRODUCTION

LLMs use natural language processing (NLP), and often artificial intelligence (AI), to process and generate text for translation, information retrieval, conversational interactions, and summarisation. Countering violent extremism (CVE) efforts are also making use of LLMs, by automatically generating counter-speech (interventions that directly address, or offer alternatives to extremist content), with the aim of dissuading individuals from extremist narratives. Scholars have argued that LLMs may assist practitioners in identifying hate speech, creating targeted counter-speech content, and engaging with users to deliver counter-speech. This article outlines the potential for LLMs as a vehicle for counter-speech, whilst highlighting the challenges that must be overcome.

POTENTIAL BENEFITS

Using LLMs to automatically generate counter-speech can assist with the increasing demand for counter-speech online, as they can be used to supplement the design and dissemination of counter-speech and may be more easily scalable than solely human-operated interventions. When identifying where counter-speech is needed, LLMs could automatically detect hate speech. Counter-speakers have also shown interest in using AI to gather relevant information to assist them in designing counter-speech content. Using LLMs to assist counter-speakers in identifying hate speech and constructing counter-speech responses could be a beneficial time-saving strategy.

It is important to reduce the strain on counter-speakers, as they can become the victims of hate speech themselves and experience negative impacts on their personal well-being. Since counter-speech requires a lot of time and effort and often involves addressing upsetting content, some counter-speakers have reported feeling overwhelmed and experiencing negative impacts on their mental health. As such, counter-speakers have suggested that it could be beneficial to use AI for counter-speech campaigns to protect their personal well-being, as it can offer them anonymity and reduce the amount of harmful content that they must identify and review, helping them to become more emotionally detached.

Counter-speakers have suggested that it could be beneficial to use AI for counter-speech campaigns to protect their personal well-being.

POTENTIAL CHALLENGES

Whilst using LLMs for counter-speech can reduce the physical and mental load on counter-speakers, there are important practical limitations and ethical challenges that should be considered.

There are questions surrounding the functionality of LLMs and their ability to produce relevant and credible counter-speech to a variety of audiences. Generating counter-speech requires large amounts of data, language comprehension, emotional intelligence, and contextual knowledge, which can be challenging to achieve with machine learning alone. As such, some LLM-generated counter-speech has been found to include factual, or grammatical errors, which may limit the effectiveness of the counter-speech within a CVE campaign.

Counter-speakers have expressed concerns that a lack of human involvement during the design and dissemination of counter-speech may have impacts on the authenticity and credibility of the counter-speech. For instance, disclosing that counter-speech was written by AI has been found to significantly reduce users perceived trust in the counter-speech. Alternatively, if it is not disclosed when AI has been used for counter-speech, transparency concerns can arise from users being unaware that LLMs are involved. It is important that LLM-generated counter-speech is manually reviewed to ensure it is authentic, accurate, and appropriate for the target audience.

Using fully automated design and delivery methods for counter-speech raises a number of ethical considerations including user privacy and bias. Privacy is a key ethical consideration for LLMs, which can gather highly personal information about individuals. For example, when conversing with a chatbot, individuals may not understand that their conversation can be read by real people. When designing counter-speech campaigns that use LLMs, user privacy should be carefully considered and protected as much as possible.

Biases can be built into LLMs by the developers, from the datasets used to train them, or during real-world implementation. Biases can result in harmful and inaccurate information being shared, which can encourage discrimination. For example, a translation LLM on Facebook incorrectly translated "good morning" written in Arabic to "hurt them" and "attack them", resulting in a Palestinian individual being arrested by Israeli authorities. Diverse groups of individuals need to be involved in the design and dissemination of LLMs to ensure that biases are mitigated and LLMs are not generating harmful or incorrect content.

THE EFFECTIVENESS OF LLM-GENERATED COUNTER-SPEECH

There is limited empirical research that assesses the effectiveness of LLM-generated counter-speech as part of a CVE campaign. Most studies focus on the ability of LLMs to generate counter-speech, leaving measures of effectiveness mostly restricted to evaluations of the counter-speech's written quality. There are, however, a smaller number of studies that do assess the practical effectiveness and impact of LLM-generated counter-speech.

Chung et al. (2021) tested an NLP tool that aimed to assist NGO practitioners in countering Islamophobia on Twitter in English, French, and Italian. The tool detected Islamophobia and then automatically composed counter-speech responses. The practitioners who tested the tool had positive feedback and felt it was an innovative addition to counter-speech writing. Importantly, some of the practitioners emphasised that the tool should not entirely replace manual writing, but instead should be used to assist them in counter-speech writing, as modifications to the generated counter-speech were often necessary. The tool was still considered to be useful as writing counter-speech from scratch reportedly took longer than modifying the counter-speech that the tool generated.

Bilewicz et al. (2021) assessed whether counter-speech messages that were generated and delivered by a bot (disguised as a real male user) could be effective at reducing verbal aggression within two subreddits. When verbal aggression was detected, the bot delivered counter-speech by directly replying to the user who posted the aggressive content. User comments posted 60 days before and after the intervention were analysed and a control group of users from other subreddits, who were not targeted with the bot, were used for comparison. Users displayed a lower proportion of verbal aggression after the intervention than before, whereas the proportion of verbal aggression in the control group remained largely the same throughout. This study suggests that bots could be used as part of counter-speech interventions.

Bär et al. (2024) explored whether LLM-generated counter-speech was effective in reducing hate on Twitter (X). Compared to manually-generated counter-speech, the LLM-generated counter-speech was found to be less effective in reducing online hate. The LLM-generated counter-speech was also associated with an increase in hate posts after users viewed the counter-speech. The researchers suggest that users may have recognised that the counter-speech was LLM-generated, causing them to react negatively. These findings highlight the potential counter-productive backfire effects that can arise from the use of LLMs to generate counter-speech.

CONCLUSION

LLMs can generate counter-speech in response to a range of extremist content, which can offer some protections to counter-speakers and may reduce the human resources that are needed to design and deliver counter-speech campaigns. However, there are some important limitations associated with the use of LLMs to generate counter-speech that must be considered. The potentially limited functionality of LLMs can result in the production of counter-speech that contains inaccuracies and the design and use of LLMs raises user privacy and bias concerns that can have harmful implications.

The small number of studies that assess the practical effectiveness of LLM-generated counter-speech offer mixed findings. LLMs show some promise in assisting practitioners with counter-speech writing and may potentially help to reduce verbal aggression online. It is important to consider that using LLMs to generate counter-speech may backfire and create increased hostility. Any use of LLMs within counter-speech campaigns needs to be accompanied by rigorous evaluation, risk assessment, and human oversight to ensure that the counter-speech being generated is relevant, factual, and non-harmful.

Using fully automated design and delivery methods for counter-speech raises a number of ethical considerations including user privacy and bias.

Ellie Rogers is a PhD candidate at Swansea University within the Cyber Threats Research Centre (CYTREC). Her research focuses on the algorithmic amplification of counter-speech as a response to online extremism.

> I am a better psychologist, a more informed applied practitioner, and the base knowledge of all of our customers is significantly higher than it was pre-CREST.
>
> – Government practitioner

CREST CELEBRATES 10 YEARS

A decade of commissioning behavioural science, collaborating across sectors, and communicating insights that have informed and shaped security practice.

> CREST's openness to multidisciplinary research is not simply rhetoric – it's influence and reach has changed behaviour and cognition.
>
> – Professor Julie Gore

CREST 10 YEAR TIMELINE

2015

CREST

2016

CREST LAUNCHES

Sunni and Shi'a Islam
(Kim Knott)
Our most popular resource ever, with over 67k downloads and 225k views, still topping monthly charts nine years on.

CREST Conference – London
The inaugural conference showcased cutting-edge research from CREST PhDs, early career researchers, five core programmes, and commissioned projects, bringing stakeholders together to explore the full breadth of CREST's work.

"...talks were pitched at exactly the right level."
– Government

2017

Sikh Activism in Britain
(Jasjit Singh)
Explores the narratives and issues that drive Sikhs in Britain to engage in visible activism, becoming a key resource for understanding Sikh radicalisation.

"Jas Singh's work on Sikh radicalisation was not seen as 'policy critical' at commissioning, but quickly became of interest and is now the go-to resource"
– Impact Assessment

"Prof Vrij's research has informed UK Aviation and Border Security interrogation policy."
– Practitioner

Model Statement Technique
(Aldert Vrij)
A simple technique for eliciting more information from cooperative interviewees, now integrated into security training and policy.

CREST Conference – Reading
The second conference focused on building the CREST community and sharing research across core and commissioned projects. Presentations featured work from ECRs alongside senior researchers.

2018

Russian Influence Campaigns
(Martin Innes)
Revealed systematic disinformation after the 2017 UK terror attacks, leading to international funding and an evidence review on social media's role in crises.

"Amazing mix of academics, policy makers, and researchers"
– Practitioner

BASS – Lancaster
The first public BASS featured 50+ talks across three themes, with keynote speakers from Canada and Australia. Over 100 delegates attended to share research and explore international perspectives.

2019

Muslims & Islam in the UK
(Kim Knott)
A set of short guides created to make CREST's in-depth research on Islam more accessible to practitioners. Used in training for nearly 1,000 end users across multiple countries and agencies.

Organisational Change
(Rosalind Searle & Charis Rice)
Widely used by managers and adopted by NHS Scotland and CPNI (now NPSA), these tools helped organisations manage change effectively.

CRESTFest – London
A two-day showcase designed for government stakeholders. ECRs presented new research and joined informal discussions, creating direct opportunities for policy–research exchange.

2020

Deradicalisation Programmes
(Sarah Marsden)
Explored approaches to disengagement and evaluation. Informed Home Office policy and led to Marsden chairing an independent advisory committee.

Understanding Conspiracy Theories
(Karen Douglas et al.)
CREST's most cited publication, this seminal 2019 paper shaped the development of conspiracy theory research as a field. It continues to guide academics, policymakers, and practitioners in understanding the psychology and social dynamics of conspiracy beliefs.

2021

Community Reporting on Terrorism
(Michelle Grossman & Paul Thomas)
Explored why families and friends hesitate to report loved ones' extremist activity, becoming a central element of delivering the UK PREVENT strategy, influencing a number of workstreams led by the national PREVENT team.

The Eliciting Information Framework
(Anna Leslie)
A practical tool for interviewing and debriefing, rapidly deployed during the pandemic.

"...today shows how much I rely on CREST in this respect."
– Practitioner

"...one of the best pieces of training in my 40 years of service."— NCA

2022

Dynamics of Violence Escalation
(Joel Busher et al.)
Explored how anti-minority and far-right activism escalates or is contained during 'hot periods'. The project and briefings received detailed praise from security practitioners.

CREST/CPD Negotiation event – Manchester
Delivered with the UK CT Negotiation Team, this private event explored communication, engagement, and influence in negotiations. 5/7 presentations came from CREST.

BASS – Lancaster
After the pandemic pause, BASS returned with 65 talks on risk, eliciting information, and disruption. It hosted 180+ delegates from eight countries, plus DSTL-funded panels on social bias.

"A great opportunity to present my work and get feedback from brilliant people." – PhD researcher

2023

BASS – Bath
BASS23 brought 59 talks and 230 delegates from seven countries. Keynotes came from King's College London, Cardiff, and DSTG Australia, alongside popular returning themes.

"...one of the most organised (and best) conferences I've been to... exceptional quality of speakers."
– Practitioner

Lethal Subcultures & the Extreme Far-Right
(Ben Lee)
Explained emerging extreme far-right movements, siege culture and lethal subcultures, helping to inform Counter Terrorism Policing training.

2024

Emergency Services Interoperability
(Nicola Power et al.)
Exposed the "principle-to-implementation" gap in emergency response, underpinning new research and a UKRI Future Leaders Fellowship.

Incel-Related Violence
(Lewys Brace)
Attracted huge public interest, surging to 40k views in one month after the Netflix TV show Adolescence came out.

The Role of Social Media in Far-Right Violence
(Olivia Brown)
After the 2024 UK riots, Brown's work informed public understanding through outlets such as Financial Times, Policing Insight, BBC, LBC, and CBC News, shaping debate on online harms and safety

BASS – St Andrews
BASS24 hosted 65+ talks, 192 delegates, 49 universities, and 85 government reps. Keynotes from Montreal and OFCOM, plus a UK policing panel, highlighted practice-driven impact.

"It was fabulous... amazing opportunities to meet up, network and exchange ideas"
– Researcher

2025

NABS+ Community Event – Crewe
Funded by UKRI, supported by CREST, NABS+ was launched a cross-sector network on behavioural and social science for security. The inaugural event gathered 29 institutions to tackle online harms, disinformation, and AI.

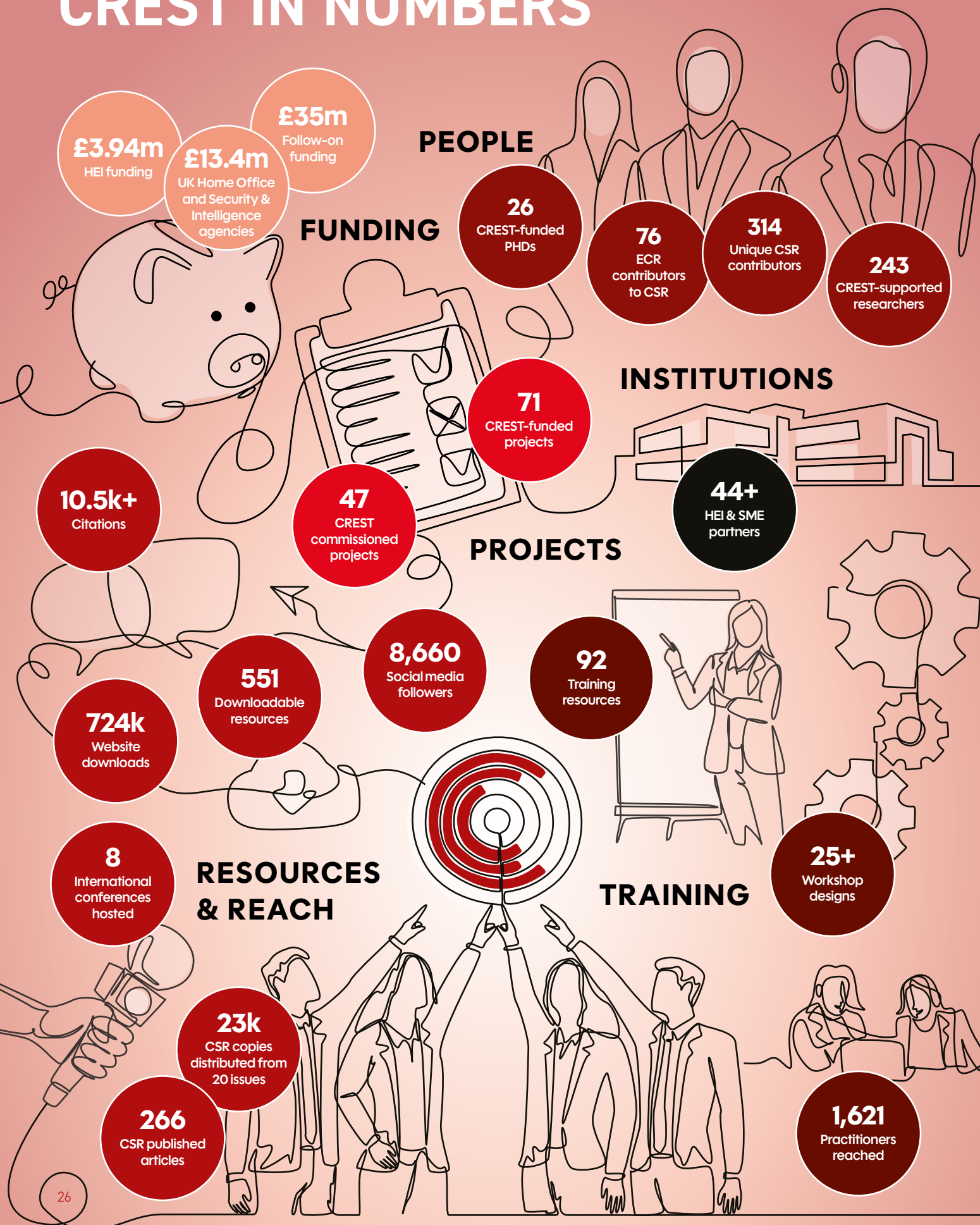
Time-Critical Questioning Framework
(Lorraine Hope et al.)
A pioneering protocol tested in live hostage-taking exercises, improving rapid information extraction under pressure.

"We've witnessed first-hand the efficacy of this pioneering technique." – Practitioner

EVENTS OUTPUTS

To read more, visit crestresearch.ac.uk/resources/10-years-of-crest

CREST IN NUMBERS



FUNDING

£3.94m
HEI funding

£13.4m
UK Home Office and Security & Intelligence agencies

£35m
Follow-on funding

PEOPLE

26
CREST-funded PHDs

76
ECR contributors to CSR

314
Unique CSR contributors

243
CREST-supported researchers

INSTITUTIONS

44+
HEI & SME partners

PROJECTS

71
CREST-funded projects

47
CREST commissioned projects

10.5k+
Citations

8,660
Social media followers

551
Downloadable resources

724k
Website downloads

TRAINING

92
Training resources

25+
Workshop designs

RESOURCES & REACH

8
International conferences hosted

23k
CSR copies distributed from 20 issues

266
CSR published articles

1,621
Practitioners reached

HILARY PILKINGTON

WHY STUDY NON-RADICALISATION?

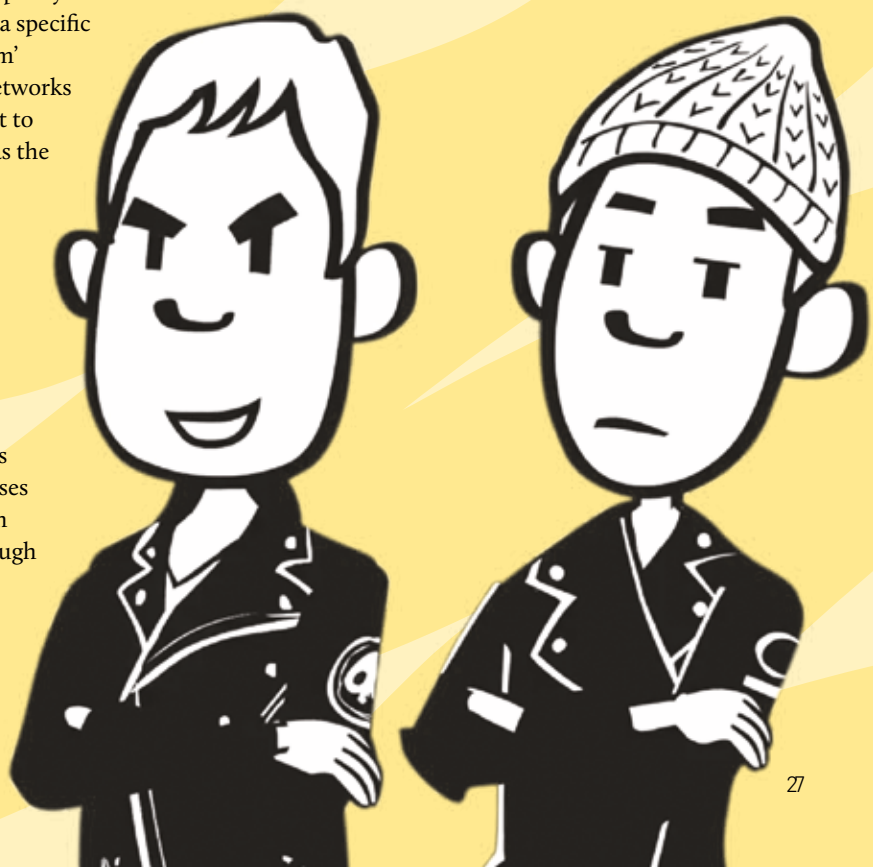
This article presents a critical, empirically grounded contribution to our understanding of radicalisation drawing on research findings presented in the book Resisting Radicalisation? Understanding Young People's Journeys through Radicalising Milieus.

The DARE (Dialogue about Radicalisation and Equality) project brought together researchers from 12 countries across Europe and its near neighbours to better understand how young people navigate radical milieus, specifically ideological and political environments associated with 'Islamist' or 'extreme-right' views. The findings led to the development of an empirically grounded critique of radicalisation.

The concept of radicalisation underpins much of the preventative strand of Counter Terrorism policy and infrastructure worldwide. In the UK, for example, this can be found in the PREVENT duty (2015 –). Revised guidance 2023 states a core objective to be to 'intervene early to support people susceptible to radicalisation' and mentions 'radicalisation', 'radicalisers' or 'radicalised' almost 200 times. Given that the concept of radicalisation was catapulted into common policy and public parlance relatively recently - in response to a specific security threat around so-called 'home-grown terrorism' following 9/11 and the bombings of public transport networks in Madrid in 2004 and London in 2005 - it is important to keep a critical eye on whether it retains its usefulness as the nature of threats to society evolve.

This article argues that a focus on understanding the 'how' of terrorism, which the concept of radicalisation offers, remains important. However, paradoxically, a focus on the process by which people move towards violent extremism has privileged the endpoint in determining what we understand as radicalisation and stunted a more holistic understanding of how people engage in, but also resist, political violence. This is because studies of radicalisation start by selecting cases where acts of terrorism or violent extremism have been committed and retrospectively identify the stages through which individuals progress towards that point. For policymakers, there is a self-evident public safety logic in focusing on actors who cross the violence threshold. For researchers, it also provides a clear target population for study and allows trajectories to that point to be traced, compared, and modelled.

Paradoxically, a focus on the process by which people move towards violent extremism has privileged the endpoint in determining what we understand as radicalisation.



However, there are inherent limitations to explaining involvement in violent extremism by studying only those who have committed terrorist acts while excluding those who move in the same milieu but do not become violent extremists. By selecting on the dependent variable (a violence outcome), the terrorist act appears de facto as the single outcome of the radicalisation process when only a small proportion of those who hold radical, or even extreme, ideas go on to commit acts of violence.

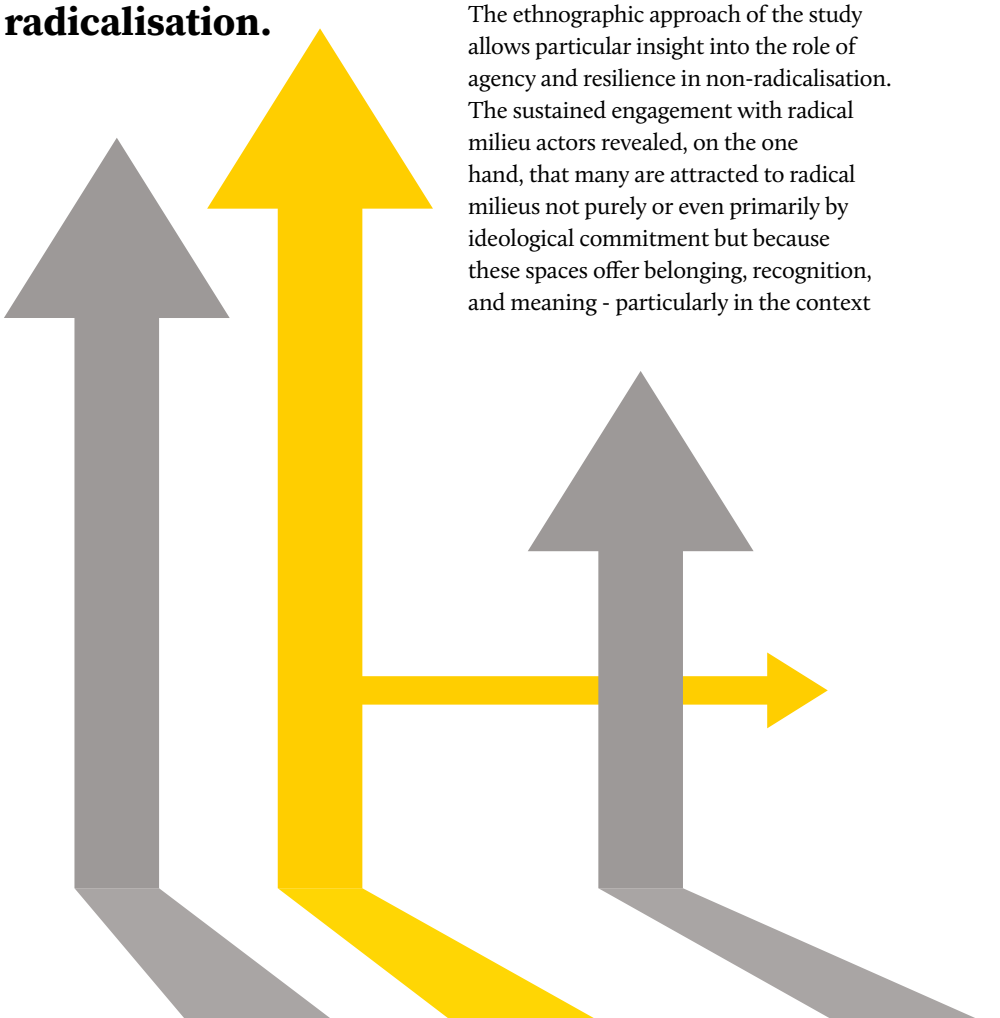
To avoid this endpoint determinism, the DARE project selected radical milieus both online and offline, as its central unit of analysis. These were milieus associated with both 'right-wing extremist' and 'Islamist' ideologies and included a wide range of contexts including political movements and parties, prisons, online forums and networks, and church and other faith groups. Researchers engaged directly with young people (aged 15-30) in situ, conducting interviews, observations, and discussions to capture lived experiences as they unfolded. In total over 500 observations and just under 400 semi-structured interviews were conducted with young people active across nine 'right-wing extremist' (184 interviews) and ten 'Islamist' (199 interviews) radical milieus.

A key finding of the study is that radicalisation should not be viewed as a linear pathway towards violence. Rather, engagement in violence laden milieus can also result in non-radicalisation. The concept of non-radicalisation is not new – a small number of studies of such pathways have sought to explain non-involvement in terrorist violence among radical actors or among wider populations in political violence laden contexts. These studies deploy distinctions between attitudinal and behavioural (or non-violent and violent) extremism to determine a non-violent 'control group' and identify the presence/absence of risk/protective factors against which trajectories ending in violence can be compared. While providing important insights, this approach has inherent limitations, not least because it relies on absolute distinctions between 'target' and 'control' groups that are themselves slippery and falls back on the determining role of a violence outcome in understanding radicalisation. It also limits the capacity to map, and explain, the range of different non-radicalisation journeys radical actors follow.

Radicalisation and non-radicalisation are viewed as existing on a continuum along which individuals engaged in radical milieus may move in either direction in trajectories of what might be called partial, stalled, or reversed radicalisation.

In contrast, the DARE project understands non-radicalisation not as the absence of radicalisation to be explained by individuals' moral opposition to violence, awareness of the costs of radical activism, or the presence of psychological, socio-demographic, or relational protective factors. Rather, radicalisation and non-radicalisation are viewed as existing on a continuum along which individuals engaged in radical milieus may move in either direction in trajectories of what might be called partial, stalled, or reversed radicalisation.

The ethnographic approach of the study allows particular insight into the role of agency and resilience in non-radicalisation. The sustained engagement with radical milieu actors revealed, on the one hand, that many are attracted to radical milieus not purely or even primarily by ideological commitment but because these spaces offer belonging, recognition, and meaning - particularly in the context

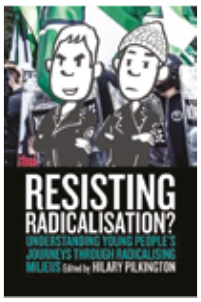


of marginalisation or social exclusion. On the other hand, it provided insight into how individuals' radicalisation journeys stall, halt, or even go into reverse as they reflect on, question, or even reject radicalising messages and agents they encounter.

While this resilience will be in part due to aforementioned individual protective factors, an important finding of our research is that milieus themselves are not just breeding grounds for extremism but multifaceted spaces in which ideas are debated, social bonds are formed, and resistance may be exercised. The research finds that radical milieus or movements can serve to inhibit violence as well as promote it and challenges the dominant assumption in policy and media discourse that exposure to radical messages inevitably increases the likelihood of radicalisation. Examples of how this plays out in both 'Islamist' and 'right-wing extremist' milieus are explored comparatively in Chapters 1 and 2 of the Resisting Radicalisation? volume and in a UK case study setting in Chapter 6. For example, a participant in the Norwegian 'right-wing extremist' milieu whose attitudinal radicalisation led to his self-identification as a Nazi and initial engagement with the Nordic Resistance Movement (NRM), halted his radicalisation journey after realising the NRM's advocation of violence was a moral 'red line' he could not cross and switched association to the non-violent Stop Islamisation of Norway (SIAN) movement. In a contrasting scenario, the expulsion of a UK 'extreme-right' milieu actor from the English Defence League (EDL), when the organisation was seeking to dissociate itself from violent behaviour, precipitated his decision to go 'full neo-Nazi' and establish his own movement focused on direct, violent action. Where people 'end' on the non-radicalisation-radicalisation continuum is shaped by the nature of the milieu as well as situational and interactional dynamics within the milieu and between milieus and wider society and institutions.

The implications for counter-extremism policy are that a strengths-based approach grounded in recognising protective factors and building resilience is likely to be more effective than targeting 'at risk' individuals for preventative intervention. The DARE research speaks particularly to the sphere of secondary prevention and has been drawn on to develop mediated dialogue tools to work with this hard to reach group, who are viewed as immune to primary intervention (because they are already certain of their positions) but who are not identified as likely beneficiaries of tertiary programmes of de-radicalisation because they neither identify themselves, nor are identified by formal agencies, as radicalised. Resilience, in this context, is understood as a capacity not just to avoid extremist involvement, but to navigate adversity, make informed choices and find belonging without resort to violence. Indeed, since preventative intervention is often experienced as 'suspectification' or 'silencing', feeling misrecognised or unfairly targeted by counter-terrorism policies can itself fuel resentment. In this way, securitised approaches to radicalisation risk exacerbating the very conditions they aim to prevent.

Where people 'end' on the non-radicalisation-radicalisation continuum is shaped by the nature of the milieu as well as situational and interactional dynamics within the milieu and between milieu and wider society and institutions.



Hilary Pilkington is Professor of Sociology at the University of Manchester, Her research focuses on young people, political activism and extremism and she was Coordinator of the Horizon 2020 DARE project. She has served as independent Commissioner to the GMCA's Preventing Violent Extremism and Promoting Social Cohesion Commission (2017-18) and to the UK Government's Independent Review into Civil Unrest in Leicester in 2022 (2023-25).

ACKNOWLEDGEMENTS

The DARE project (*https://sites.manchester.ac.uk/dare*) received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No.725349 and was coordinated by the author. Publications reflect only the views of the authors; the European Commission and Research Executive Agency are not responsible for any information they contain.

NICK CROSSLEY

SOCIAL NETWORKS AND RESILIENCE

What makes a social network less resilient? It depends on the dangers it faces, and the people within it, with these factors creating trade-offs rather than clear-cut answers.

Networks have structures, and their structural properties and variations can be tested for vulnerabilities. But unlike physical systems, social networks are made up of people whose own behaviours, perceptions, and choices introduce additional layers of complexity.

SOCIAL NETWORKS

A network has two basic elements: a set of nodes and a set of ties connecting particular pairs of those nodes. Anything, potentially, might count as a node or tie if it is meaningful to define it as such for research purposes. In specifically social networks, nodes are often either human actors or organisations, and ties are some kind of relationship or transaction between them. Together, ties and nodes form a structure which has measurable properties, and which can be visualised as a graph (see Figure One).

NETWORK RESILIENCE

Network resilience is usually defined by reference to the likelihood that removing a small proportion of nodes and/or ties would significantly alter a network's structure, thereby preventing it from delivering the desired outcomes. If removal of a few nodes would fragment a network into disconnected 'components' and this would significantly disrupt the flow of desired goods through it, for example, then the network would be said to lack resilience. In the abstract this makes sense but when we add details of concrete examples, particularly involving human beings, things get more complicated.

CENTRALISATION VS DECENTRALISATION

Resilience depends on network flow and threat type. For example, if a virus is spreading, rapid fragmentation into disconnected 'components' may actually help protect many nodes and ties. While this would break the integrity of the network as a whole, it could allow parts to continue functioning. In this scenario, a network that fragments easily might be considered more resilient than one that remains fully connected, since full connectivity would expose every node to infection.

Studies of network resilience often focus upon 'random attacks'. Researchers remove nodes and/or ties randomly and consider the impact upon a network's structure and functioning. This is often inappropriate for social networks, particularly in a national security context where attacks are frequently targeted.

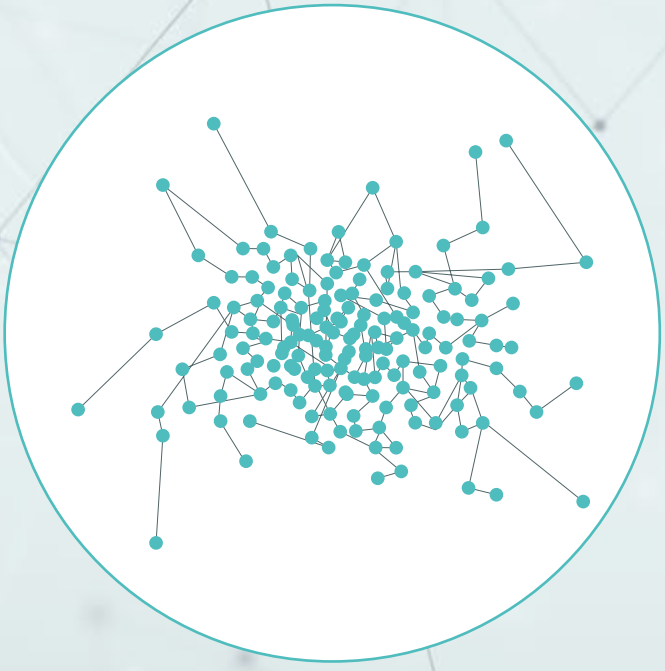


Figure One: A Basic Network Graph

Decentralisation comes with trade-offs: while it may reduce vulnerability to attack, it can also slow communication and hinder coordination, introducing different kinds of weaknesses.

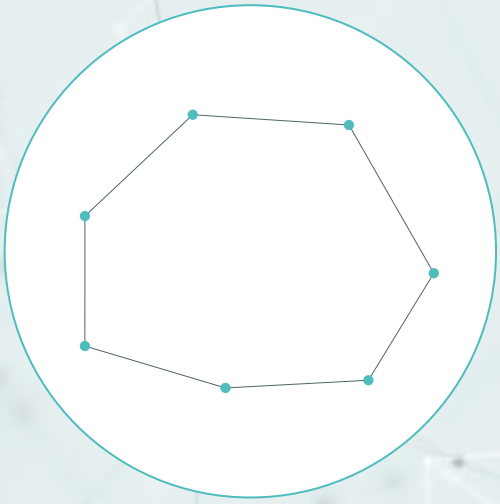
To give an illustration, in his popular book, Linked, Barabási argues that 'scale-free networks', that is, very large networks involving a tiny number of 'hubs' which account for a disproportionately large number of connections, are highly resilient. Hubs hold scale-free networks together, indirectly connecting the many other nodes which connect to them. Their loss would have a major negative impact on the structure and functioning of the network. However, because they form such a small minority of nodes within a scale-free network (perhaps one in several million), the probability of one of them being removed in a random process is vanishingly small. Random attacks are much more likely to affect other nodes, which have little significance for the structure and functioning of the network.

This is good news if chance is our main adversary but not if we are dealing with sophisticated cyber-terrorists who will attack hubs deliberately. (Barabási's subsequent work acknowledges some such complexities).

Scale-free networks are highly centralised, built around a small number of nodes with disproportionately many connections (see Figure Two for an idealised model). By contrast, more decentralised networks (where nodes have a roughly similar number of ties) may be more resilient to targeted attacks, as they lack obvious single points of failure (Figure Two). This resilience is often cited in the case of malign COM networks such as 764, and it also underpinned the shift in some Islamic terror networks towards a model of "leaderless jihad". However, decentralisation comes with trade-offs: while it may reduce vulnerability to attack, it can also slow communication and hinder coordination, introducing different kinds of weaknesses.

A Completely Centralised Network

Random damage is more likely to affect one of the peripheral nodes than the central node but the central node is an obvious target for a deliberate attack.



A Completely Decentralised Network

There is no obvious target for an attacker, but communication is likely to be slow and coordination therefore difficult.

Figure Two: Centralisation and Decentralisation

THE ROLE OF NETWORK DENSITY

Increasing network density, that is, increasing the proportion of all pairs of nodes which are connected, would potentially increase speed of communication without necessarily making a network any more centralised. It would also increase the number of indirect paths of connection between nodes, reducing their dependence upon any one path and making connections more resilient to damage. However, higher density requires nodes to maintain more connections which can be costly and which increases node exposure to dangers within the network. In addition, it allows disinformation, viruses, and other 'bads' to spread more effectively through a network, again reducing resilience.

As a final point, in contrast to the physical systems which are often the focus in discussions of resilience, the human nodes of social networks typically perceive and attach meaning to 'attacks'; repairing damage, addressing the source of their vulnerability, and rounding upon their attacker. Removing key players from a terrorist network is not the same as removing a fuse from a circuit board. Activity may be frustrated in the short-term but not stopped.

Nick Crossley is a Professor of Sociology at Manchester University and is a co-founder and co-director of the Mitchell Centre for Social Network Analysis.

CLAIRE HAZELDEN

COMMUNITY RESILIENCE IN NORTHERN IRELAND:
A FRAMEWORK FOR TRANSFORMATION

Community resilience is "the ability of a community to prepare, respond and recover from disruptive challenges". In Northern Ireland (NI), where the legacy of the 'Troubles' still shapes daily life, resilience must extend beyond recovery to transformation.

The Executive Programme on Paramilitarism and Organised Crime (EPPOC) has developed a Community Resilience Framework to support communities most affected by deprivation, coercive control, and criminal exploitation. This article explores how EPPOC's framework enables transformative change, offers insights into resilience in post-conflict societies, and highlights lessons for wider security and peacebuilding efforts.

UNDERSTANDING THE CONTEXT: PARAMILITARISM AS MULTI-LAYERED HARMS

Paramilitarism costs NI in the region of £750 million annually and affects up to 40% of the population. While some paramilitary groups engage in community activity, others are involved in criminality, organised crime, and terrorism. Most are no longer cohesive or militaristic in structure.

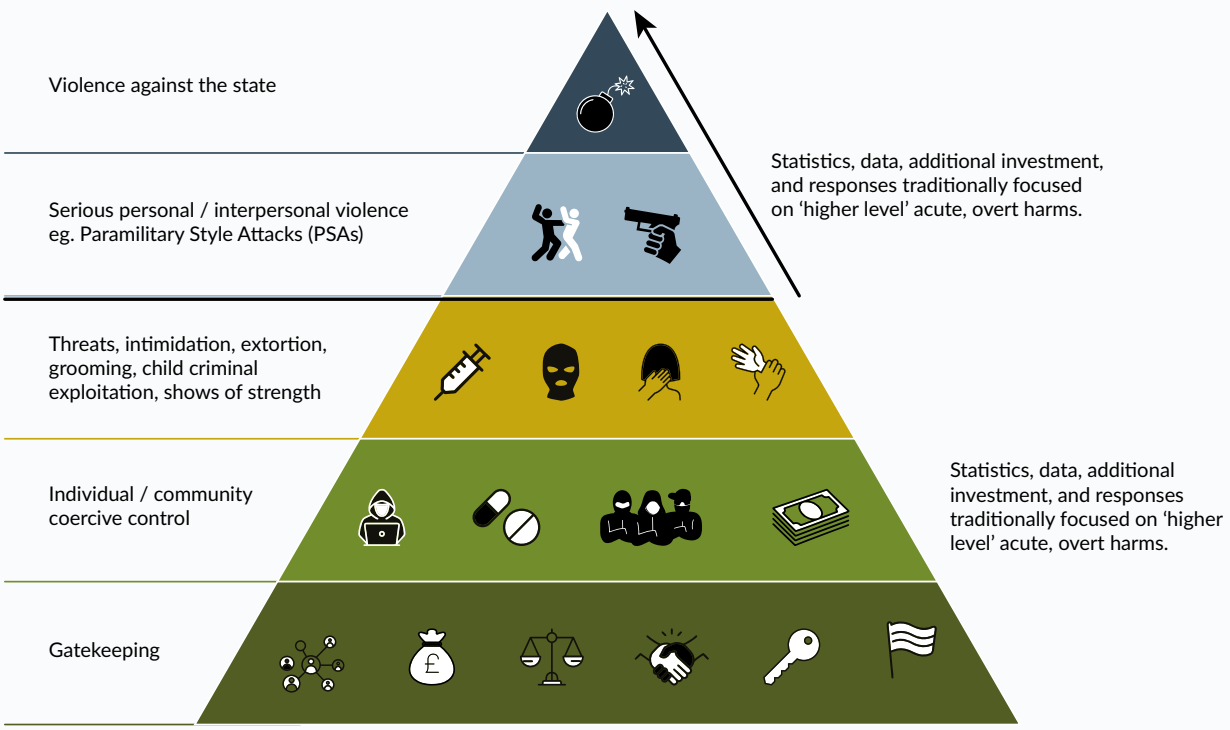


Figure 1: Typology of Paramilitary Harm

As a small, post-conflict, and devolved nation, NI faces unique complexities. Its transition from violent ethnic conflict has been globally recognised but has been neither smooth nor easy. The mere presence of armed non-state actors represents a persistent threat in areas most affected.

Figure One outlines the range of harms caused by paramilitaries, from visible and well documented to less obvious but equally damaging. EPPOC aims to address all of these, blending international evidence-based practice with innovative local approaches.

EPPOC research also highlights the profound impact of Adverse Childhood Experiences (ACEs) across NI . Sixty percent of adults reported at least one ACE, and nearly 20% experienced four or more. These levels are strongly associated with poor mental and physical health, educational disruption, and increased vulnerability to harm in adulthood. Almost a third of respondents (some born after the Good Friday Agreement) experienced conflict-specific adversities, including paramilitary threats and collective violence. These were most prevalent in deprived communities, compounding existing inequalities and reinforcing cycles of trauma.

EPPOC

Established in 2016, EPPOC delivers interventions across Government Departments, law enforcement agencies, local councils, and community and voluntary sector partners to tackle paramilitarism and its wider harms.

It operates across NI, with projects varying in scale, focus, and methodology based on assessed needs. All projects must demonstrate added value and contribute to shared, measurable medium and long-term objectives.

EPPOC uses a public health model to understand root causes, identify effective interventions, and scale them for maximum benefit. This involves a cross-sector, phased approach to intervention (see Figure 2), functioning like a relay team to drive lasting change.

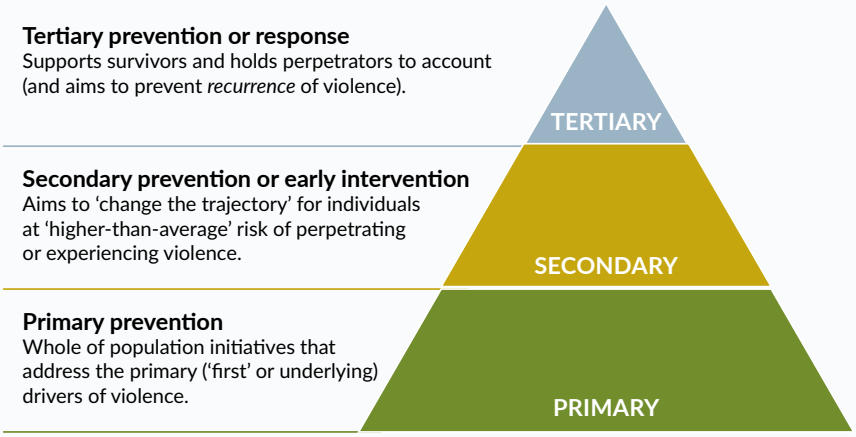


Figure 2: Public health Approach - stages of intervention

DEFINING RESILIENCE IN A POST-CONFLICT SETTING

The need to address the root causes of paramilitarism in NI is well-documented. EPPOC's approach combines immediate threat reduction, early intervention, and long-term structural change, targeting inequalities and intergenerational trauma. Equally important is recognising the strengths within communities that are so powerful in driving change.

The Community Resilience Framework considers three capacities:

- **Absorptive resilience** - the ability to withstand shocks and "keep going"
- **Adaptive resilience** - is capacity to adjust, cope, or "work around" challenges
- **Transformative resilience** - the capacity to reshape systems when ecological, economic, or social structures are no longer tenable

While absorptive and adaptive resilience help communities survive, transformative resilience enables them to thrive.

HOW CAN COMMUNITY RESILIENCE SUPPORT TRANSFORMATIONAL CHANGE?

EPPOC defines community resilience across seven measurable domains:

1. Strengthening relationships and engagement to drive collective action
2. Improving access to existing support services
3. Building local skills and knowledge
4. Enhancing infrastructure
5. Amplifying marginalised community voices
6. Fostering hope and belief in the power to create change
7. Improving PSNI-community relations

A wide range of projects, from youth engagement and restorative practices to community development and support for ex-prisoners, contribute to each domain. Evaluation is led by Queens University Belfast through an ongoing Action Research model and project-level reporting, enabling adaptive learning across EPPOC.

INSIGHTS FROM DELIVERY: WHAT'S WORKING

Since 2019, Communities in Transition (CiT) has placed local people at the heart of decision-making, empowering communities to identify challenges and lead solutions. In eight areas, community-led forums shaped action plans on key issues like youth services, housing, and community safety. This trust-based, collaborative approach reflects the idea that change is best achieved through collective action grounded in personal relationships. It has delivered concrete outcomes - from reimaging paramilitary murals to creating community safety forums across NI.

One of the most powerful indicators of transformative change is the restoration of hope.

Thousands have engaged in sustained activity as part of CiT, with tens of thousands attending community events such as health fairs, cultural exhibitions, and social action projects. These figures reflect a growing sense of agency in communities historically marked by distrust and disengagement.

EPPOC has invested heavily in skills development. The Diploma in Community Development Practice equips individuals to lead local initiatives, with many participants going on to volunteer, work, and lead future iterations. By developing the confidence and expertise to challenge coercive control and advocate for their communities, EPPOC is seeding long-term structural change. Similarly, the Developing Women in the Community project has increased women's leadership from 35.3% to 60.8%, and volunteering from 46.4% to 79.2%, an important shift given the gendered nature of exclusion and inequality.

New community organisations have been established, and existing ones strengthened. Many report increased capacity to address local vulnerabilities and work effectively with agencies.

Through CiT, thousands have accessed support services, including mentoring, tailored plans, and drop-in clinics. Multi-agency collaboration has enabled cross-referral between projects and statutory providers. Community surveys and events have addressed local concerns, ensuring responsive action.

One of the most powerful indicators of transformative change is the restoration of hope. In the Developing Women in the Community project, measures of belonging rose from 67.3% to 77.7%, and a sense of agency in shaping their areas rose significantly. These shifts directly counter the control narratives exploited by paramilitary groups.

EPPOC's impact is not uniform across all areas. Persistent disparities remain, and addressing structural imbalances will require sustained investment and political will. The foundations laid by EPPOC offers a model for how resilience can be built from the ground up, through participation, capacity-building, and the reweaving of social ties.

WHAT CAN THIS TELL US ABOUT RESILIENCE IN POST CONFLICT SOCIETIES?

EPPOC's work offers valuable insights for building resilience in post-conflict societies:

- **Beyond traditional measures:** Resilience must move past frameworks focused solely on social capital and services to include psychological and intergenerational effects of trauma
- **Inequality and trauma:** High levels of ACEs in deprived areas show that resilience cannot be divorced from structural inequality and ongoing harm.
- **A dynamic process:** Resilience is a dynamic process of recovery, adaptation, and resistance
- **Trauma-informed and place-based:** Systems must build protective factors, foster benevolent experiences, and disrupt cycles of adversity
- **Visible and invisible harm:** Resilience strategies must address both overt violence and subtle, cumulative harms
- **Context matters:** Absorbative, adaptive, and transformative resilience manifest differently and must be contextually understood
- **Strength from within:** Community strengths such as solidarity, pride, and creativity are key drivers of change
- **Local empowerment:** Empowering communities to lead decision-making fosters autonomy, legitimacy, ownership, and sustainable development
- **Restoring hope:** Transformative resilience requires healing, and rebuilding belief in the possibility of change is vital for lasting peace

TRANSFORMATIVE RESILIENCE: EPPOC'S APPROACH AND LESSONS FOR POST-CONFLICT SOCIETIES

Resilience is not about enduring adversity indefinitely; it's about gaining the strength to overcome it. EPPOC applies a public health approach to violence: treating violence like a disease means not only alleviating symptoms but transforming the conditions that allow harm to recur. This includes building economic and political alternatives, restoring hope and driving systemic change.

Communities that organise to anonymously report crime or steer teenagers away from paramilitary gangs demonstrate resilience that confronts adversity instead of withstanding it. This includes leveraging strength into efforts that break the grip of violence and deprivation, rather than reinforcing the status quo.

EPPOC views resilience as an evolving process shaped by relationships, resources and power. Its work responds to local contexts, builds on strengths, and challenges the conditions that sustain paramilitary influence. It repositions communities from passive recipients of support to active agents of change.

Developing meaningful and lasting community resilience takes time, collaboration, innovation, and courage. EPPOC's framework, grounded in targeted interventions, robust evaluation, and community participation, charts a path toward an inclusive, peaceful, and confident NI. A recent NI Affairs Committee report welcomed the "trauma-informed and public health approach" being taken to address the systemic causes of paramilitarism. This reflects a growing consensus that investing in long-term community capacity is essential for an enduring peace.

NI's communities continue to shoulder the unacceptable burdens of serious violence, trauma and social division. These harms constrain potential and drive unsustainable demand on services. Other post-conflict regions face similar challenges, including the presence of armed groups, the emergence of organised crime, and lack of confidence in the police and wider justice systems. EPPOC offers lessons worth sharing.

> **Resilience is not about enduring adversity indefinitely; it's about gaining the strength to overcome it.**

KEY LESSONS FROM EPPOC'S VIOLENCE REDUCTION WORK

1 **Data driven collaboration:** Evidence-based approaches must be operationalised to encourage ownership among all partners

2 **Shared, staged outcomes:** Agreeing long term goals with manageable short-term outcomes can build consensus on sensitive topics

3 **Measure what matters:** All outcomes should be measurable, and addressing data gaps is essential to track progress

4 **Value the unknowns:** Understanding what isn't known is as important as understanding what we do

5 **Cross-sector translation:** Shared goals help bridge gaps between government, academia, and the community and voluntary sector

6 **Delivery matters:** What we do is important but so is how we do it. Identifying key enablers and shared behaviours is key

7 **Sustainable commitment:** Systemic change demands sustainable funding and policy commitment.

There is a clear moral imperative for prioritising this work and also a compelling economic one.

- Home Office data shows public health approaches prevent an estimated 243 violence without injury offences per 100,000 people, with a £4.10 return per £1 invested
- The Cardiff Model for Violence Prevention showed £82 benefit for every £1 spent, including £14.80 for the health service and £19.10 for the justice system
- EPPOC estimates a reduction of 112 paramilitary attacks saving £1.57-£2.23m and 84 fewer incidents of violence with injury in 2022, saving £1.18m. These figures only reflect paramilitary related harm; scaling to wider harms could generate even greater savings.

EPPOC's work illustrates that resilience is not about accepting the world as it is but reshaping it. By empowering communities, building capacity, and investing in long-term change, NI shows that even after conflict, communities can transform their resilience into a foundation for lasting security and hope.

Claire Hazelden is a Research Analyst for the NI Executive Programme on Paramilitarism and Organised Crime (EPPOC).

CHLOE L SQUIRES

THE NEXT GENERATION OF RESILIENCE

The co-ordinator of the Terrorism and Political Violence Next Generation Network, based at the Handa Centre for the Study of Terrorism and Political Violence (CSTPV) at the University of St Andrews, reflects on the benefits that NextGen networks offer to the field.

INTRODUCTION

Next Generation Networks ('NextGen networks') are becoming increasingly popular in the field of counterterrorism, defence, and security. A NextGen network is an organisation that aims to support early career researchers (ECRs) navigate academia and the broader research environment.

In Summer 2024, the Terrorism and Political Violence Next Generation Network was founded at CSTPV, University of St Andrews. Since its establishment, the network has cultivated a community of over 300 members from around the world. Each month, members receive a newsletter featuring opportunities for ECRs worldwide, including funding calls, paper submissions, and online webinars. The network also hosts online events tailored to the needs outlined by the community. These have included sessions on publishing, ethical, and fieldwork challenges, and undertaking research in complex and changing political contexts.

HOW CAN NEXTGEN NETWORKS SUPPORT RESILIENCE?

Within a supportive environment, NextGen networks promote resilience across the field by enabling international and interdisciplinary connections, facilitating knowledge exchange, and identifying emerging challenges. For scholars working on terrorism and political violence, the opportunities afforded by NextGen networks support the development of skills that promote robust and insightful research and analysis.

INTERDISCIPLINARY PROBLEM SOLVING

Bringing together ECRs helps build communities of emerging researchers with diverse interdisciplinary and international networks; an asset for individual researchers and research teams alike. These connections encourage transversal approaches to understanding and countering complex and emerging issues. Peer engagement fosters awareness of current and emerging research agendas and supports collaborative thinking, relationships between practitioners and academics, and grant writing.

SUPPORT MECHANISMS

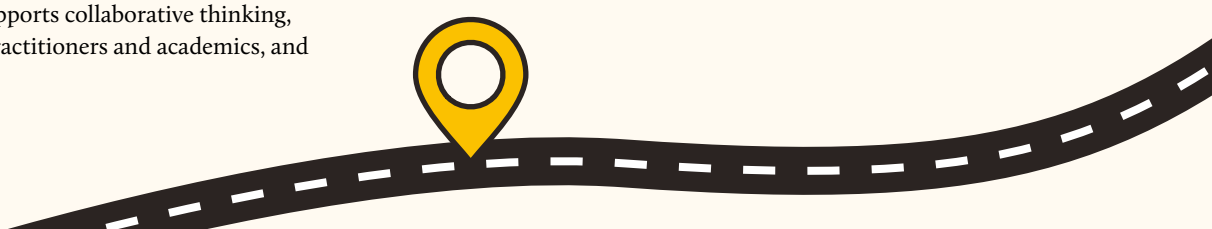
NextGen networks can offer support for ECRs and PhD researchers navigating complex research environments which can be emotionally challenging and, at times, isolating. Due to the interdisciplinary nature of terrorism studies, many researchers are based in departments such as Psychology, Criminology, Law, or Politics, where topic support may be limited. NextGen networks can be responsive to these needs and provide support beyond the parameters of an individual's university.

For example, VOX-pol provides extensive resources aiming to support online extremism and terrorism researchers to protect their emotional and mental wellbeing, and has its own NextGen network for those working on online extremism. NextGen initiatives provide spaces for ECRs to connect with peers tackling similar challenges and can offer guidance on navigating challenging research environments.

UNDERSTANDING THE FIELD

Understanding practitioner and policymaker interests and approaches can be a challenge for many ECRs and PhDs. Though their work is often highly relevant, because ECRs and PhDs often do not have direct access to these communities, they can lack insight into how these environments operate, or how their work can be applied. Some networks have aimed to address these gaps and support ECR access to policy and practice.

For example, RUSI NextGen delivered an online session to provide insight into the operation of a Think Tank. More recently, the NABS+ network launched a mentoring scheme, developed to support ECRs 'build their skills, knowledge and confidence through one-to-one conversations with an experienced mentor'. Support offered to ECRs and PhDs through these efforts helps to deepen their understanding of the field, enhance the insights that can contribute to policy and practice, as well as building their confidence in doing so.



ECR networks equip emerging scholars and practitioners with broad skills, connections, and understanding.

INTERNATIONAL CONNECTIONS AND KNOWLEDGE EXCHANGE

In the early stages of their careers, PhD students and ECRs commonly bring new perspectives and innovative research methodologies to the field. The focus on contemporary issues and new solutions make postgraduate conferences particularly valuable for knowledge exchange by providing opportunities to showcase new insights from international contexts.

In 2025, the TPV NextGen network hosted the 9th Society for Terrorism Research Postgraduate conference, bringing together students from Europe, North America, and Asia. A direct result of the NextGen network, the conference facilitated new international connections while providing the opportunity to share knowledge and develop a deeper understanding of current and emerging research agendas. The supportive and collegiate atmosphere fostered in this environment enables ECRs to develop public speaking skills and refine academic arguments, building resilience in the communication of their work to diverse audiences.

THE FUTURE OF NEXTGEN

Early career researchers face a range of challenges, from academic challenges, navigating a complex research landscape, building professional connections with peers and institutions, and protecting their emotional wellbeing. NextGen networks offer practical solutions to the challenges ECRs face navigating the field by equipping emerging scholars and practitioners with broad skills, connections, and understanding which cannot be accessed in many universities and research institutions.

Fostering supportive communities has a direct impact on the capacity of ECRs. The work of networks such as VOX-Pol Next Gen, RUSI Next Gen, NABS+ (even those which do not directly focus on the 'NextGen') are vital to developing new research agendas and prioritising researcher wellbeing.

Ultimately, NextGen networks contribute to building the next generation of resilient research- and resilient researchers.

Chloe Squires is a NABS+ Research Associate. She also co-ordinates the Terrorism and Political Violence NextGen network based at the Handa Centre for the Study of Terrorism and Political Violence, University of St Andrews.

SÖREN HENRICH

ILLUMINATING THE FORENSIC BLACK BOX:
UNDERSTANDING RADICALISATION IN SECURE SETTINGS

Over the last six years in UK secure settings, care teams have asked whether knowledge of radicalisation in the community applies to forensic institutions such as prisons or secure psychiatric hospitals. In response, I conducted a systematic review of the literature up to April 2023, examining psychological mechanisms of radicalisation across different settings.

This article highlights key themes, research gaps, and areas for further investigation from the 96 reviewed studies, which followed the PRISMA standards. This included transparent inclusion and exclusion criteria for search strings, outlined databases and quality appraisal.

We know less than we think about radicalisation in forensic settings. Theoretically, two processes can occur:

1 Individuals with no prior history of extremism become radicalised for the first time in the setting; or

2 Radicalised individuals admitted to a forensic facility continue their extremism pathway, or their radicalisation is significantly enhanced.

Of the 96 studies, only five studies from the review exclusively examined psychological dynamics within these environments. Study quality, when assessed in the review for all 96 articles, showed only 56% of studies were rated as good, while the rest were rated fair or poor. For the studies pertaining to forensic settings, three were rated good and two were fair.

While clinical experts possess valuable knowledge concerning the context set by forensic settings, academic research access remains difficult, leading some scholars to describe prisons as potential "breeding grounds for terrorists". Since then, findings have relativised this notion. While some individuals radicalise behind bars, evidence suggests prison settings can also disrupt radicalisation. Rehabilitation efforts may be more effective than previously thought, though empirical data remains limited. Even less is understood about radicalisation in forensic psychiatric hospitals, and more broadly, the role of mental health in radicalisation remains contentious, with studies both supporting and rejecting a connection.

OVERLAPPING RISK FACTORS

A central finding from the review was that risk factors for violent extremism significantly overlap with those for general violence. This raises a key question: how do we disentangle those two risk dynamics in risk assessments for an already complex forensic population?

Key risk factors included:

- **Criminal history & violent tendencies:** Identified in 39 of the 96 studies, prior offences were commonly linked with extremist violence, suggesting shared underlying factors. By definition, everyone in forensic settings has a criminal history, thus posing a unique challenge for clinicians to identify those particularly vulnerable to radicalisation.
- **Group processes:** Social influences, such as peer networks reinforcing violent ideations, increase radicalisation risk (36 studies). Rigid group norms and violence-endorsing shared narratives, which deepen divisions between groups, are particularly pertinent in forensic settings, in my clinical experience.
- **Strain & adversity:** Discrimination, marginalisation, and personal grievances heighten susceptibility to radicalisation (29 studies). In forensic populations, perceived strain is naturally high due to involuntary detainment, making this a highly relevant dynamic to consider.

THE ROLE OF IDEOLOGY: OVERESTIMATED?

Extremist attitudes were the most researched topic (41 of 96 studies), yet findings were mixed. Some authors argued ideology drives violent extremism, while others saw it as a post-hoc justification for actions motivated by other factors. The latter 'post hoc explanation' is a common occurrence also for individuals who commit non-extremist violence and are then placed in forensic settings, making the presence of violent ideation per se not only a central facet of risk assessments but also a pivotal factor to address in forensic interventions to reduce aggression.

MENTAL HEALTH: A MURKY PICTURE

Findings from 31 studies remain inconclusive. Some suggest links between conditions like depression, anxiety, and personality

disorders with violent extremism, while others find no clear association. Based on my own clinical work experience, it is not the mere presence of a diagnosis and associated symptoms that matters, but whether these symptoms can be plausibly related to extremist behaviours and violence risk (i.e., explaining the relevance of the symptoms is more important than mere presence, as explored in our multi-study that followed this review). A comprehensive examination of how mental health symptoms and violent extremism may interact can be found in '*Violent Extremism: A Primer for Mental Health Practitioners*', edited by Caroline Logan.

Despite guidance from experts like Al-Attar, uncertainty can still challenge forensic practitioners, especially in how to establish clear care pathways when mental health symptoms and extremism-related attitudes and behaviours coexist. Without stronger evidence, determining best practices remains difficult.

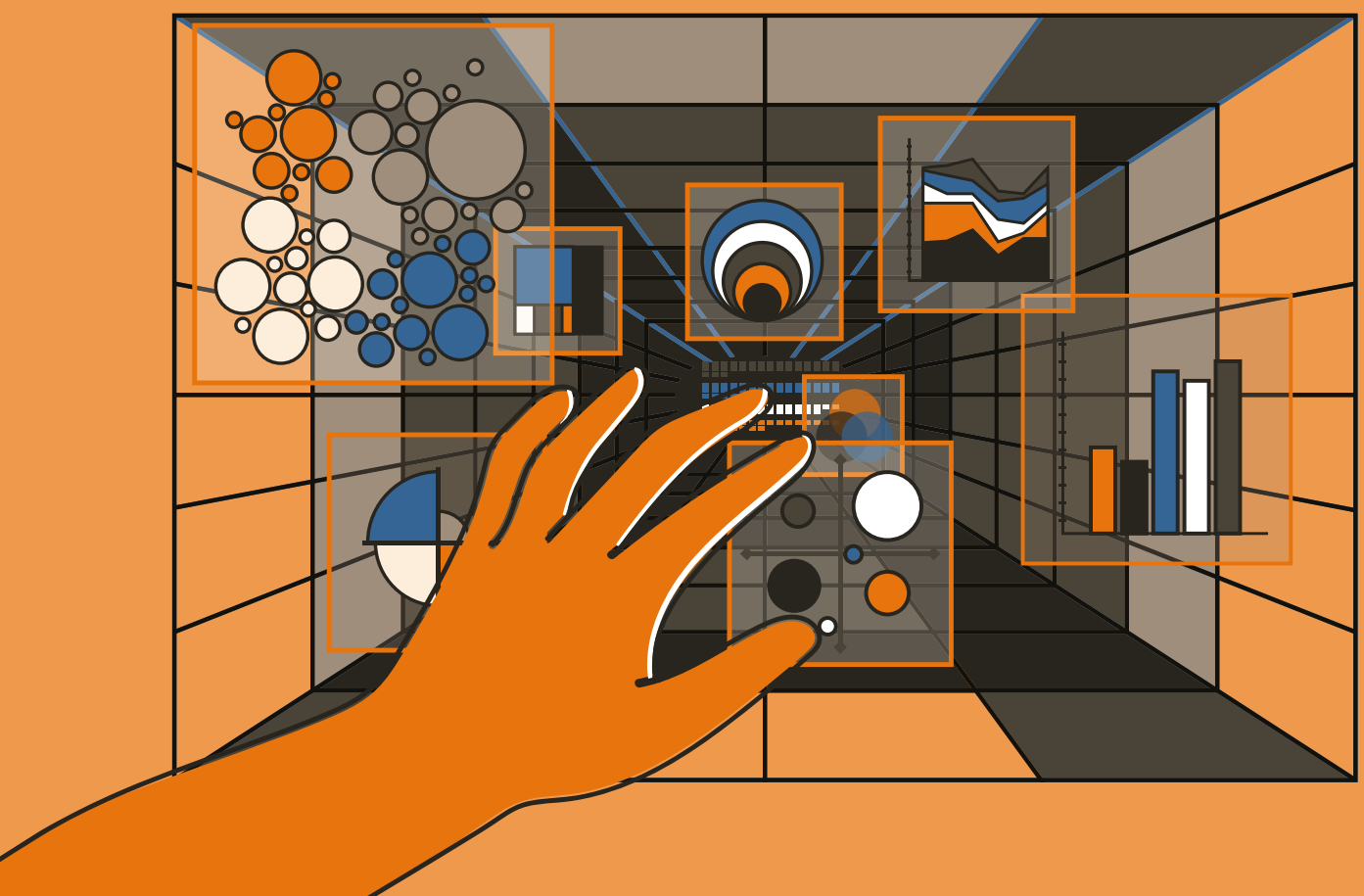
WHAT'S MISSING?

Knowledge gaps identified in the radicalisation review included:

1. **Forensic populations are understudied:** Research overwhelmingly focuses on community samples, leaving a critical gap in understanding radicalisation in secure settings and caring for those individuals.
2. **Protective factors are overlooked:** Studies emphasise risk factors, but little is known about mitigating influences like resilience, social support, and adaptive coping. Protective factors are likely to be highly relevant in forensic care settings.
3. **Lack of primary data:** Many studies rely on publicly available information rather than direct engagement with radicalised individuals in forensic settings. However, this is not necessarily due to researchers' lack of trying but because access is restricted due to national security concerns and ethical barriers. More qualitative research, including interviews, is essential.
4. **Mostly Western perspective:** Only 14 out of 93 articles had an explicitly non-Western focus. With both international actors and national forensic populations becoming more diverse, the broadening of research input beyond the US and UK is crucial.

Stronger collaboration between researchers and forensic practitioners is key to addressing these challenges and gaps. From the perspective of working in both research and mental health settings, I would suggest there is need for clearer ways to break down the complex behaviours and dynamics encountered in forensic settings, making it easier to understand how radicalisation overlaps with other types of violence. This approach would help understand mental health issues with more clarity and apply insights from community research to forensic settings while identifying emerging challenges across disciplines.

Dr Sören Henrich is a lecturer and has published papers in Forensic Psychology at the Manchester Metropolitan University. He delivers international training on risk and threat assessment and remains a clinical advisor for secure forensic settings in the UK.



EMILY QUIN

THE COPPER'S NOSE

SPIDEY SENSES AREN'T JUST FOR SPIDER-MAN

It's a phrase that echoes across canteens, briefing rooms, custody suites. In policing, this isn't just a saying; it's a phenomenon known as having 'The Copper's Nose'.

"SOMETHING JUST DIDN'T FEEL RIGHT..."

Have you ever walked into a situation and instantly known that something just wasn't right? For most of us, these unsettling moments are thankfully rare, but for police officers and frontline personnel, the ability to detect and respond to the unexpected can be the difference between life and death. This phenomenon is known in policing as 'The Copper's Nose' (TCN).

My PhD research, grounded in the lived experience of police officers, set out to understand what the phenomenon means in practice: how officers develop this sense, how it's used, and what it reveals about modern policing. Far from being mystical or magical, TCN is a finely honed cognitive and psychophysiological mechanism that aids in an officer's perception, interpretation, and response to abnormalities and risks, often in complex, ambiguous, and fast-moving environments. While many may call it an instinct, TCN acts as a context-dependent mediator between an officer and their immediate environment to facilitate a way of 'knowing' that's constantly evolving, sharpened through experience, exposure, and adaptability.

> **Officers actively learn what's 'normal' so they can more easily detect the abnormal.**

MORE THAN A HUNCH

Greater than a hunch – which could be described as belief formation – officers I interviewed described TCN as being instinctive and visceral in nature.

"As soon as that kicks in physically, I'm properly getting ready for like a fight or flight kind of situation..."

In these moments, an officer might not yet know exactly what's wrong, but they know something is. One officer explained it as a "gut instinct" and another likened it to "something activating", throwing the mind into hyper-awareness where attention narrows.

"Your heart speeds up, your thought process speeds up, and you start waking up a little bit."

This is the perceiving and attending element of TCN. It's what primes officers to notice out-of-place stimuli in the environment. Often pre-verbal and difficult to articulate, it's immediately felt and prompts the engagement of greater conscious, cognitive processing. This state of heightened awareness is embedded within each of us, but is often informed by what officers have experienced before – both on the job, and off of it – honing their skillset further.

PATTERN RECOGNITION UNDER PRESSURE

Much of this phenomenon revolves around what officers refer to as baseline deviations where something 'doesn't quite fit'. Rather than being surprised by novelty, officers actively learn what's 'normal' so they can more easily detect the abnormal. Best understood as a form of rapid interpretation, officers draw from accumulated knowledge to quickly assess environments and determine what might require further attention. However, TCN isn't infallible. Officers recognise that it can, at times, "lead [them] down the garden path", and acknowledge that acting on a gut-feeling alone doesn't guarantee a good result. What's crucial in developing a 'good' copper's nose is how these instincts are refined over time, embracing feedback, reflective practice, and most importantly, a willingness to be wrong.

FROM 'INTUITION' TO EXPERIENCE

One of the most striking themes was how this sense develops. Officers recurrently spoke about the impact of life experience before policing, and the importance of learning on the job. Critically, this kind of experience is impossible to replicate in the classroom, with many officers stating the importance of working alongside experience, reputable officers.

"For me, to watch them and observe how they do things...if I learn even one skill or one method of dealing with something, that's me tooled up for the next time..."

This updating process is both individually and socially influenced. Officers build exposure through experience and hone it through engagement, practice, and the wisdom of others. Importantly, developing the copper's nose requires more than time on the job; it demands professional curiosity.

KNOWING WHERE, WHEN, AND WHO

The research also explored how officers navigate their environments. Contributing to self-efficacy, TCN shapes how officers position themselves physically and professionally. Some officers described their nose as being their principle threat-detection mechanism. For others, their nose serves as a decision-making tool to determine when a threshold to act has been reached. Fundamentally, it's a critical part of the discretionary skill set, assisting officers to act not just according to procedure, but according to situational judgement. But this discretion isn't always celebrated. Some officers felt there exists organisational discomfort around TCN, stating that "too many freakishly good copper's noses" could generate inconvenient outcomes where resources aren't available to manage an increase in criminal detection.

IMPLICATIONS FOR POLICING AND THE FRONTLINE

While The Copper's Nose may not be something easily taught in a classroom, organisations can create the right conditions for it to develop. This means field-based training, encouraging proactive practice, and creating space for officers to learn from experienced colleagues with the appropriate practice oversight in place.

The culmination of this research also now offers a grounded theoretical framework, informed by scientific literature and primary data, that makes 'The Copper's Nose' more than just an anecdote. It presents opportunities for empirical testing and refinement, offering policing a chance to better understand, support, and develop this valuable capability.

> **For police officers and frontline personnel, the ability to detect and respond to the unexpected can be the difference between life and death.**

Dr Emily Quin, PhD, CPsychol, is a Postdoctoral Researcher at the University of Cambridge, where she specialises in applied psychology across 'frontline' settings. She obtained her PhD at the Institute of Criminology, Cambridge, where she generated a substantive theory of The Copper's Nose (policing 'sixth sense').

READ MORE

Read more about some of the research that our contributors mention in their articles. We've flagged up those that are open access and given links to online versions where they are available. For full references and citations please visit crestresearch.ac.uk/magazine/system-resilience

NICK CROSSLEY
SOCIAL NETWORKS AND RESILIENCE

Barabási, L. (2002). Linked, New York, Basic Books.

Crossley, N., Edwards, G., Harries, E., & Stevenson, R. (2012). Covert social movement networks and the secrecy-efficiency trade-off: The case of the UK suffragettes (1906–1914). *Social Networks, 34*(4), 634–644. *https://doi.org/10.1016/j.socnet.2012.07.004*

Sageman, M. (2011) Leaderless Jihad, Philadelphia, University of Pennsylvania Press. *https://doi.org/10.2307/j.ctt3fhbht*

Scott, J. (2017). Social Network Analysis, London, Sage.

KYLE S. H. DOBSON
TRANSPARENCY STATEMENTS RESTORE TRUST & RESILIENCY

Dobson, K. S., Dittmann, A. G., & Yeager, D. S. (2025). A transparency statement improves trust in community-police interactions. *Nature Communications, 16*(1), 2285. *https://doi.org/10.1038/s41467-024-55709-6*

Dube, O., MacArthur, S. J., & Shah, A. K. (2025). *A cognitive view of policing. Quarterly Journal of Economics, 140*(1), 745-791. *https://doi.org/10.1093/qje/qjae039*

Olivett, V. J., & March, D. S. (2024). The civilian's dilemma: Civilians exhibit automatic defensive responses to the police. *Journal of Personality and Social Psychology, 126*(5), 841-851. *https://doi.org/10.1037/pspi0000439*

Tapp, S. N., & Davis, E. J. (2022, November). *Contacts between police and the public, 2020* (BJS Special Report No. NCJ 304527). U.S. Department of Justice, Bureau of Justice Statistics. *https://bjs.ojp.gov/library/publications/contacts-between-police-and-public-2020*

GRETA ESPOSITO
INTERPRETING AND TRANSLATION IN THE POLICE SETTING: HUMAN VS AI

Baraldi, C., & Gavioli, L. (2013). Mediating patient-centred interaction: The case of doctors' questions. *Salute e Società, 12*(1), 104–120.

Berk-Seligson, S. (1990). *The bilingual courtroom: Court interpreters in the judicial process*. University of Chicago Press.

College of Policing. (2020). *Advice for forces on using interpretation and translation services*. *https://www.police.uk/pu/policing-in-the-uk/national-police-language-services-npls/advice/*

Filipović, L. (2022). The tale of two countries: Police interpreting in the UK vs in the US. *Interpreting, 24*(2), 254–278. *https://doi.org/10.1075/intp.00080.fil*

Gallai, F. (2013). *Understanding discourse markers in interpreter-mediated police interviews* (Doctoral dissertation). University of Salford.

Hale, S. (2002). How faithfully do court interpreters render the style of non-English speaking witnesses' testimonies? A data-based study of Spanish–English bilingual proceedings. *Discourse Studies, 4*(1), 25–47. *https://doi.org/10.1177/1461445602004001020l*

Hale, S. (2007). *Community interpreting*. Springer

Hale, S. (2014). The treatment of register variation in court interpreting. *The Translator, 20*(1), 39–54. *https://doi.org/10.1080/13556509.1997.10798987*

Krouglov, A. (1999). Police interpreting: Politeness and sociocultural context. *The Translator, 5*(2), 285–302. *https://doi.org/10.1080/13556509.1999.10799045*

Loakes, D. (2024). Automatic speech recognition and the transcription of indistinct forensic audio: How do the new generation of systems fare? *Frontiers in Communication, 9*, 1281407. *https://doi.org/10.3389/fcomm.2024.1281407*

Nakane, I. (2014). *Interpreter-mediated police interviews: A discourse-pragmatic approach*. Springer. *https://doi.org/10.1057/9781137443199*

Wadensjö, C. (1992). *Interpreting as interaction: On dialogue-interpreting in immigration hearings and medical encounters* (Linköping studies in arts and science, No. 83). Linköping University. *https://doi.org/10.7146/hjlcb.v7i12.25067*

CLAIRE HAZELDEN
COMMUNITY RESILIENCE IN NORTHERN IRELAND: A FRAMEWORK FOR TRANSFORMATION

The EPPOC Community Resilience framework is based on work undertaken by Lucy Geddes (CIT Benefits Handbook & Resilience workshop report) from Co-operation Ireland, who utilised a co-design approach to identify indicators of community resilience to Paramilitarism, criminality and organised crime, drawing heavily on Interpeace's Resilience for Peace framework.

Cabinet Office. (2019, June). *Community Resilience Development Framework: A reference tool for the delivery of strategic approaches to community resilience development, at the Local Resilience Forum level in collaboration with non-statutory partners*. *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/828813/20190902-Community_Resilience_Development_Framework_Final.pdf*

Cardiff University. *The Cardiff model for violence prevention.* Cardiff University. *https://www.cardiff.ac.uk/documents/2665796-the-cardiff-model-for-violence-prevention*

The Centre for Collaboration in Community Connectedness. *https://c4.ac.uk*

Colm, W., Bunting, L., Davidson, G., Doherty, N., McCartan, C., Mulholland, C., & Shevlin, M. (2025, February). The Prevalence and Impact of Adverse Childhood Experiences in Northern Ireland. The Executive Programme on Paramilitarism and Organism Crime (EPPOC). *https://www.endingtheharm.com/wp-content/uploads/2025/02/Impact-of-Adverse-Childhood-Experiences-report-060225-1.pdf*

Communities in Transition. *https://www.executiveoffice-ni.gov.uk/articles/communities-transition-background*

Developing Women in the Community. *https://www.communities-ni.gov.uk/articles/developing-women-community-programme-phase-three-2025-2027-information-and-guidance-notes*

Diploma in Community Development Practice Diploma in Community Development Practice 2025 - 2026. *https://www.ruralcommunitynetwork.org/eventbrite-event/diploma-in-community-development-practice-2025-2026*

Ending the Harm. *https://www.endingtheharm.com*

House of Commons, Northern Ireland Affairs Committee. (2024, January 23). The effect of paramilitary activity and organised crime on society in Northern Ireland. *https://committees.parliament.uk/publications/43124/documents/214542/default*

HM Government (2019, June). Community Resilience Development Framework. Cabinet Office. *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/828813/20190902-Community_Resilience_Development_Framework_Final.pdf*

Interpeace. (2016). *Assessing resilience for peace: A guidance note* [PDF file]. *https://www.interpeace.org/frameworks-for-assessing-resilience-far/*

Northern Ireland Affairs Committee. (2024, February 1). *The effect of paramilitary activity and organised crime on society in Northern Ireland [Summary]*. UK Parliament. *https://publications.parliament.uk/pa/cm5804/cmselect/cmniaf/43/summary.html*

Queen's University Belfast (2024, 10 December) [web article]. Queen's part of new £10 million project to build stronger communities across the UK. *https://www.qub.ac.uk/News/Allnews/2024/Queenspartofnew10millionprojecttobuildstrongercommunitiesacrosstheUK.html*

Walsh, C. (2023) Developing Women in the Community Programme Evaluation. Queen's University Belfast. *https://pureadmin.qub.ac.uk/ws/files/526195926/DFC_report_2023.pdf*

SÖREN HENRICH
ILLUMINATING THE FORENSIC BLACK BOX: UNDERSTANDING RADICALISATION IN SECURE SETTINGS

Our review and following research:

Henrich, S., Ireland, J. L., & Lewis, M. (2025). Radicalisation across the community and forensic units: A systematic literature review on the psychology of violent extremism. In T. C. Silva & M. Kordaczuk-Wąs (Eds.), *Preventing and countering violent extremism and radicalisation: Evidence-based policy and practice* (pp. 163–215). Policy Press. *https://doi.org/10.51952/9781447370949.ch005*

Henrich, S., Ireland, J. L., Lewis, M., & Ireland, C. A. (2025). The eco-system of extremist violence (ES-EV): Exploration of radicalisation in forensic psychiatric populations. *The Journal of Forensic Practice, Advance online publication*. *https://doi.org/10.1108/JFP-12-2024-0076*

Mental Health and Radicalisation:

Al-Attar, Z. (2021). Severe mental disorder and terrorism: When psychosis, PTSD and addictions become a vulnerability. In *Violent extremism* (pp. 52–72). Routledge.

Corner, E., & Gill, P. (2020). Psychological distress, terrorist involvement and disengagement from terrorism: A sequence analysis approach. *Journal of Quantitative Criminology, 36*, 499–526. *https://doi.org/10.1007/s10940-019-09420-1*

Understanding Risk in Conjunction with Protective Factors:

Ireland, J. L., Levtova, Y., Abi Semaan, C. M., Steene, L. M., Henrich, S., Gaylor, L., & Chu, S. (2024). Risk and protective factors in risk assessment: Predicting inpatient aggression in adult males detained in a forensic mental health setting. *Aggressive Behavior, 50*(3), e22150. *https://doi.org/10.1002/ab.22150*

Prison and Radicalisation:

Mulcahy, E., Merrington, S., & Bell, P. J. (2013). The radicalisation of prison inmates: A review of the literature on recruitment, religion and prisoner vulnerability. *Journal of Human Security, 9*(1), 4–14. *https://doi.org/10.12924/johs2013.09010004*

Silke, A., Morrison, J., Maiberg, H., Slay, C., & Stewart, R. (2021). The Phoenix Model of disengagement and deradicalisation from terrorism and violent extremism. *Monatsschrift für Kriminologie und Strafrechtsreform, 104*(3), 310–320. *https://doi.org/10.1515/mks-2021-0128*

HOPE MCVEAN
SYNTHETIC CELEBS & FABRICATED FAMILY: PROTECTING AGAINST A NEW BREED OF CYBERCRIME

Brewster, T. (2021, October 14). Fraudsters Cloned Company Director's Voice In $35 Million Heist, Police Find. Forbes. *https://www.forbes.com/sites/thomasbrewster/2021/10/14/huge-bank-fraud-uses-deep-fake-voice-tech-to-steal-millions/*

Bristow, T. (2023, October 9). Keir Starmer suffers UK politics' first deepfake moment. It won't be the last. Politico. *https://www.politico.eu/article/uk-keir-starmer-labour-party-deepfake-ai-politics-elections*

Bunn, A. (2023, May 15). Artificial Imposters—Cybercriminals Turn to AI Voice Cloning for a New Breed of Scam. McAfee. *https://www.mcafee.com/blogs/privacy-identity-protection/artificial-imposters-cybercriminals-turn-to-ai-voice-cloning-for-a-new-breed-of-scam/*

Khan, A., Malik, K. M., Ryan, J., & Saravanan, M. (2023). Battling voice spoofing: A review, comparative analysis, and generalizability evaluation of state-of-the-art voice spoofing countermeasures. *Artificial Intelligence Review, 56*(1), 513–566. *https://doi.org/10.1007/s10462-023-10539-8*

McVean, H. (2024). *Synthetic celebs and deepfake deception: Discrimination between human and AI-generated speech of known vs unknown speakers* (Master's dissertation, Lancaster University). Lancaster University Repository.

Meaker, M. (2023, October 3). Slovakia's Election Deepfakes Show AI Is a Danger to Democracy. Wired. *https://www.wired.com/story/slovakias-election-deepfakes-show-ai-is-a-danger-to-democracy*

Mlot, S. (2023, January 31). People Are Still Terrible: AI Voice-Cloning Tool Misused for Deepfake Celeb Clips. PCMag. *https://uk.pcmag.com/news/145199/people-are-still-terrible-ai-voice-cloning-tool-misused-for-deepfake-celeb-clips*

Rose, R. & Cohen, M. (2024, May 23). Political consultant behind fake Biden AI robocall faces charges in New Hampshire. CNN. *https://edition.cnn.com/2024/05/23/politics/new-hampshire-ai-robocall-biden-charges/index.html*

Stupp, C. (2019, August 30). Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case. Wall Street Journal. *https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402*

Verma, P. (2023, March 5). They thought loved ones were calling for help. It was an AI scam. The Washington Post. *https://www.washingtonpost.com/technology/2023/03/05/ai-voice-scam*

ANNE PETERSCHECK
ARTIFICIAL INTELLIGENCE AND GENDERED SECURITY

Albahar, M., & Almalki, J. (2019). Deepfakes: Threats and countermeasures systematic review. *Journal of Theoretical and Applied Information Technology, 97*(22), 3242–3250. *https://www.jatit.org/volumes/Vol97No22/7Vol97No22.pdf*

Berg, J., Kamiński, K., Konopczyński, F., Ładna, A., Rosłaniec, K., & Troszyński, M. (2025). Generative AI and jobs: A refined global index of occupational exposure (Working Paper No. 140). International Labour Organization. *https://www.ilo.org/publications/generative-ai-and-jobs-refined-global-index-occupational-exposure*

Cadore, C. (2020). *Why women's economic rights are the key to reducing violence against women (and how experiences of informal workers tell us all we need to know)*. Womankind Worldwide. *https://www.womankind.org.uk/why-womens-economic-rights-are-the-key-to-reducing-violence-against-women/*

Hawkins, W., Russel, C., & Mittelstadt, B. (2025, April 17). Dramatic rise in publicly downloadable deepfakes image generators. *Oxford Internet Institute*. *https://www.oii.ox.ac.uk/news-events/dramatic-rise-in-publicly-downloadable-deepfake-image-generators/*

Laffier, J., & Rehman, A. (2023). Deepfakes and harm to women. *Journal of Digital Life and Learning, 3*(1), 1–21. *https://doi.org/10.51357/jdll.v3i1.218*

McNeilly, M., & Smith, P. (2023). *Will generative AI disproportionately affect the jobs of women? Kenan Insights*. UNC Kenan-Flagler Business School. *https://kenaninstitute.unc.edu/kenan-insight/will-generative-ai-disproportionately-affect-the-jobs-of-women*

Namvarpour, M., Pauwels, H., & Razi, A. (2025). *AI-induced sexual harassment: Investigating contextual characteristics and user reactions of sexual harassment by a companion chatbot*. arXiv:2504.04299. *https://doi.org/10.48550/arXiv.2504.04299*

Girlguiding. (2025). One in four teens have seen pornographic deepfakes online. *Girlguiding*. *https://www.girlguiding.org.uk/about-us/press-releases/one-in-four-teens-have-seen-pornographic-deepfakes-online*

Prama, T. T., Amrin, J. F., Anwar, M. M., & Sarker, I. H. (2025). *AI enabled user-specific cyberbullying severity detection with explainability*. arXiv:2503.10650. *https://doi.org/10.48550/arXiv.2503.10650*

HILARY PILKINGTON
WHY STUDY NON-RADICALISATION?

Busher, J., Holbrook, D., & Macklin, G. (2023). How the "internal brakes" on violent escalation work and fail: Toward a conceptual framework for understanding intra-group processes of restraint in militant groups. *Studies in Conflict & Terrorism, 46*(10), 1960–1983. *https://doi.org/10.1080/1057610X.2021.1872156*

Coolaset, R. (2019). Radicalisation: The origins and limits of a contested concept. In N. Fadil, M. de Koning, & F. Ragazzi (Eds.), *Radicalisation in Belgium and the Netherlands: Critical perspectives on violence and security* (pp. 29–51). I.B. Tauris. *http://hdl.handle.net/1854/LU-8617730*

Cragin, K. (2014). Resisting violent extremism: A conceptual model for non-radicalization. *Terrorism and Political Violence, 26*(2), 337–353. *https://doi.org/10.1080/09546553.2012.714820*

Cragin, K., Bradley, M., Robinson, E. & Steinberg, P. (2015). 'What factors cause youth to reject violent extremism? Results of an exploratory analysis in the West Bank. Rand Corporation. Available at: *https://www.rand.org/pubs/research_reports/RR1118.html*

Grossman, M. (2021). Resilience to violent extremism and terrorism: A multisystemic analysis. In M. Ungar (Ed.), *Multisystemic resilience: Adaptation and transformation in contexts of change* (pp. 293–317). Oxford University Press.

HM Government. (2023). Prevent duty guidance: Guidance for specified authorities in England and Wales, HMSO. *https://www.gov.uk/government/publications/prevent-duty-guidance*

Horgan, J. (2008). From profiles to pathways and roots to routes: Perspectives from psychology on radicalization into terrorism. *The Annals of the American Academy of Political and Social Science, 618*(1), 80–94. *https://doi.org/10.1177/0002716208317539*

Hussain, A., Pilkington, H., Simcock, K., Nicholas, J., Vickers, H., & Rogerson, L. (2019). *Talking our way out of conflict: Critical reflections on 'mediated dialogue' as a tool for secondary level CVE*. Commission for Countering Extremism. *https://www.gov.uk/government/publications/talking-our-way-out-of-conflict*

Jaskoski, M., Wilson, M., & Lazareno, B. (2020). Approving of but not choosing violence: Paths of nonviolent radicals. *Terrorism and Political Violence, 32*(2), 257–274. *https://doi.org/10.1080/09546553.2017.1364638*

Knight, S., Keatley, D., & Woodward, K. (2022). Comparing the different behavioral outcomes of extremism: A comparison of violent and non-violent extremists, acting alone or as part of a group. *Studies in Conflict & Terrorism, 45*(8), 682–703. *https://doi.org/10.1080/1057610X.2019.1680192*

Malthaner, S., & Waldmann, P. (2014). The radical milieu: Conceptualizing the supportive social environment of terrorist groups. *Studies in Conflict & Terrorism, 37*(12), 979–998. *https://doi.org/10.1080/1057610X.2014.962441*

Pilkington, H. (Ed.). (2023). *Resisting radicalisation? Understanding young people's journeys through radicalising milieus*. Berghahn Books. *https://www.berghahnbooks.com/title/PilkingtonResisting*

Reidy, K. (2019). Benevolent radicalization: An antidote to terrorism. *Perspectives on Terrorism, 13*(4), 1–13. *https://www.jstor.org/stable/26756699*

Schuurman, B., & Carthy, S. (2024). Understanding (non)involvement in terrorist violence: What sets extremists who use terrorist violence apart from those who do not? *Criminology & Public Policy, 23*(1), 119–152. *https://doi.org/10.1111/1745-9133.12626*

EMILY QUIN

THE COPPER'S NOSE: SPIDEY SENSES AREN'T JUST FOR SPIDER-MAN

Bechara, A. (2011). The somatic marker hypothesis and its neural basis: Using past experiences to forecast the future in decision making. In M. Bar (Ed.), *Predictions in the brain: Using our past to generate a future* (pp. 122–133). Oxford University Press. *https://doi.org/10.1093/acprof:oso/9780195395518.003.0048*

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 351*(1346), 1413–1420. *https://doi.org/10.1098/rstb.1996.0125*

Epstein, S. (2010). Demystifying intuition: What it is, what it does, and how it does it. Psychological Inquiry, 21(4), 295–312. *https://doi.org/10.1080/1047840X.2010.523875*

Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 683–703). Cambridge University Press. *https://doi.org/10.1017/CBO9780511816796.038*

Gigerenzer, G. (2023). *The intelligence of intuition*. Cambridge University Press.

Klein, G. A. (2011). *Streetlights and shadows: Searching for the keys to adaptive decision-making*. MIT Press. *https://doi.org/10.7551/mitpress/8369.001.0001*

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185(4157), 1124–1131. *https://doi.org/10.1126/science.185.4157.1124*

Willis, J. J. (2013). Improving police: What's craft got to do with it? *Ideas in American Policing, 16*, 1–13. *https://www.researchgate.net/publication/258120819_Improving_Police_What's_Craft_Got_to_Do_With_It*

Worrall, J. (2013). The police sixth sense: An observation in search of a theory. *American Journal of Criminal Justice, 38*, 306–322. *https://doi.org/10.1007/s12103-012-9176-0*

ELLIE ROGERS

AUTOMATICALLY GENERATING COUNTER-SPEECH: OPPORTUNITIES AND CHALLENGES OF USING LLMS FOR CVE

Baele, S. J, & Brace, L. (2024). *AI Extremism: Technologies, Tactics, Actors. Vox-Pol Network of Excellence*. *https://voxpol.eu/wp-content/uploads/2024/04/DCUPN0254-Vox-Pol-AI-Extremism-WEB-240424.pdf*

Bär, D., Maarouf, A. & Feuerriegel, S. (2024). Generative AI may backfire for counterspeech. *arXiv:2411.14986*. *https://doi.org/10.48550/arXiv.2411.14986*

Bilewicz, M., Tempska, P., Leliwa, G., Dowgiałło, M., Tańska, M., Urbaniak, R. & Wroczyński, M. (2021). Artificial intelligence against hate: Intervention reducing verbal aggression in the social network environment. *Aggressive Behaviour, 47*(3), 260-266. *https://doi.org/10.1002/ab.21948*

Blasiak, K., Risius, M., & Matook, S. (2021, December 1). 'Social Bots for Peace': A Dual-Process Perspective to Counter Online Extremist Messaging. *Forty-Second International Conference on Information Systems, Austin 2021*. International Conference on Information Systems, Austin, Texas. *https://www.researchgate.net/publication/354906975_Social_Bots_for_Peace_A_Dual-Process_Perspective_to_Counter_Online_Extremist_Messaging*

Briggs, R. & Feve, S. (2013). *Review of programs to counter narratives of violent extremism*. Institute for Strategic Dialogue. *https://www.publicsafety.gc.ca/lbrr/archives/cn28580-eng.pdf*

Buerger, C. (2024). *Counterspeech: A literature review*. Dangerous Speech Project. *https://www.dangerousspeech.org/libraries/counterspeech-a-literature-review*

Chung, Y. (2022). Counter Narrative Generation for Fighting Online Hate Speech [unpublished doctoral dissertation]. University of Trento. *https://iris.unitn.it/retrieve/handle/11572/338563/544707/PhD_Thesis_YiLingChung.pdf*

Chung, Y., Sinem Tekiroğlu, S., Tonelli, S. & Guerini, M. (2021). Empowering NGOs in countering online hate messages. *Online Social Networks and Media, 24*, 100150. *https://doi.org/10.1016/j.osnem.2021.100150*

Coltri, M. A. (2024). The Ethical Dilemma with Open AI ChatGPT: Is It Right or Wrong to Prohibit It? *Athens Journal of Law, 10*(1), 119-130. *https://doi.org/10.30958/ajl.10-1-6*

Davey, J., Tuck, H., and Amarasingam, A. (2019). *An imprecise science: Assessing interventions for the prevention, disengagement and de-radicalisation of left and right-wing extremists and Countering Political Extremism*. Institute for Strategic Dialogue. *https://www.isdglobal.org/wp-content/uploads/2019/11/An-imprecise-science-1.pdf*

Dennis, A. R., Kim, A., Raimi, R., and Ayabakan, S. (2020). User reactions to COVID-19 screening chatbots from reputable providers. *Journal of the American Medical Informatics Association, 27*(11), 1727–1731. *https://doi.org/10.1093/jamia/ocaa167*

Donath, J. (2020). Ethical Issues in Our Relationship with Artificial Entities. In Markus D. Dubber, Frank Pasquale, and Sunit Das (Eds), *The Oxford Handbook of Ethics of AI* (pp. 52-73). Oxford University Press. *https://doi.org/10.1093/oxfordhb/9780190067397.013.3*

Gebru, T. (2020). Race and Gender. In M.D. Dubber, F. Pasquale & S. Das (Eds.), *The Oxford Handbook of Ethics and AI* (pp. 252-269). Oxford University Press. *https://doi.org/10.1093/oxfordhb/9780190067397.013.16*

Howard, J. W. (2021). Terror, Hate and the Demands of Counter-Speech. *British Journal of Political Science, 51*(3), 924–939. *https://doi.org/10.1017/S000712341900053X*

Irfan, M., Abdulaziz, Z. & Anwar, M. (2023). Unleashing transformative potential of artificial intelligence (AI) in countering terrorism online radicalization extremism and possible recruitment. *Global Strategic & Securities Studies, 8*(4), 1-15. *http://dx.doi.org/10.31703/gsssr.2023(VIII-IV).01*

Mukherjee, A. (2023). *AI and Ethics: A computational perspective*. IOP Publishing. *https://iopscience.iop.org/book/mono/978-0-7503-6116-3*

Mun, J., Buerger, C., Liang, J. T., Garland, J. & Sap, M. (2024). Counterspeakers' Perspectives: Unveiling Barriers and AI Needs in the Fight against Online Hate. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, 742, 1-22. *https://doi.org/10.1145/3613904.3642025*

Naveed, H., Ullah Khan, A., Shi, Q., Saqib, M., Anwar, S., Usman, M., Akhtar, N., Barnes, N., and Mian, A. (2025). A Comprehensive Overview of Large Language Models. *ACM Transactions on Intelligent Systems and Technology, 16*(5), 1-72. *https://doi.org/10.1145/3744746*

Susser, D. (2020). Ethical Considerations for Digitally Targeted Public Health Interventions. *American Journal of Public Health, 110*(53), S290-S291. *https://doi.org/10.2105/AJPH.2020.305758*

Tanprasert, T., Fels, S. S., Sinnamon, L., and Yoon, D. (2024). Debate Chatbots to Facilitate Critical Thinking on YouTube: Social Identity and Conversational Style Make A Difference. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, 805, 1–24. *https://doi.org/10.1145/3613904.3642513*

Vallecillo-Rodríguez, M. E., Ráez, A. M. & Martín-Valdivia, M. T. (2023). Automatic counter-narrative generation for hate speech in Spanish. *Procesamiento del Lenguaje Natural, 71*, 227-245. *https://doi.org/10.26342/2023-71-18*

Wu, C., Wang, Y, Zhang, Y., Wang, H., and Pang, Y. (2025). Confront hate with AI: how AI-generated counter speech helps against hate speech on social Media? *Telematics and Informatics, 101*, 102304. *https://doi.org/10.1016/j.tele.2025.102304*

KT SONNEN
PLANNED RESILIENCE: ARE WE SETTING OURSELVES UP TO FAIL?

Bectu. (2025). *18-hour days, heatstroke and sleeping on the floor: Life behind the scenes for UK festival workers*. *https://bectu.org.uk/news/18-hour-days-heatstroke-and-sleeping-on-the-floor-life-behind-the-scenes-for-uk-festival-workers*

Benson, S.G. & Dundis, S.P. (2003). Understanding and motivating health care employees: Integrating Maslow's hierarchy of needs, training and technology. *Journal of Nursing Management, 11*(5), 315–320. *https://doi.org/10.1046/j.1365-2834.2003.00409.x*

CIPD. (2023). *Urgent action needed on UK job quality as new research suggests people are less engaged, less fulfilled and more stressed in their working life* (CIPD Good Work Index 2023). *https://www.cipd.org/uk/about/press-releases/good-work-index-job-quality-new-research/*

Drury, J. & Winter, G. (2003). Social Identity as a source of Strength in Mass Emergencies and Other Crowd Events. *International Journal of Mental Health, 32*(4). *https://doi.org/10.1080/00207411.2003.11449599*

Grandner, M. (2019). *Sleep and health*. Academic Press. *https://doi.org/10.1016/C2017-0-02743-X*

Health and Safety Executive. (1998). *Provision and Use of Work Equipment Regulations 1998 (PUWER) – overview*. *https://www.hse.gov.uk/work-equipment-machinery/puwer-overview.htm*

Health and Safety Executive. *Planning for incidents and emergencies*. *https://www.hse.gov.uk/event-safety/incidents-and-emergencies.htm*

Health and Safety at Work Act (1974) [specifically S2 and S3]. *https://www.legislation.gov.uk/ukpga/1974/37/contents*

National Protective Security Authority. *Motivation within the Security Industry*. *https://www.npsa.gov.uk/system/files/documents/npsa-guard-force-motivation.pdf*

Webb, E., Perry, M., & Fennelly, L.J. (2015). 18 – Employee motivation theory and application. In Davies, S.J., Hertig, C.A., & Gilbride, B.P. (Eds.), *Security supervision and management* (4th ed., pp. 231–240). Butterworth-Heinemann. *https://doi.org/10.1016/B978-0-12-800113-4.00018-3*

CHLOE L SQUIRES
THE NEXT GENERATION OF RESILIENCE

The TPVNextGen Network organises online events, distributes a monthly newsletter, and acts as a hub for support and collaboration to nurture the next generation of researchers. Established in 2024, the Terrorism and Political Violence Next Generation Network supports PhD candidates and early career researchers working on the broad scope of terrorism and political violence. *https://cstpv.wp.st-andrews.ac.uk/terrorism-and-political-violence-next-generation/*

Centre for Research and Evidence on Security Threats. *ECR Hub*. NABS+. *https://crestresearch.ac.uk/nabs/early-career-researchers-hub*

Royal United Services Institute. *NextGen*. RUSI. *https://www.rusi.org/nextgen*

VOX-Pol. (2023). *Researcher Resources*. VOX-Pol. *https://voxpol.eu/researcher-resources*

JAMES STEVENSON
TECHNOLOGY EVOLVES THE TACTICS: PREPARING FOR THE RISE OF TERRORIST AI HARMS

Borgonovo, F., Rizieri Lucini, S., & Porrino, G. (2024, February 23). *Weapons of mass hate dissemination: The use of artificial intelligence by right-wing extremists*. Global Network on Extremism & Technology. *https://gnet-research.org/2024/02/23/weapons-of-mass-hate-dissemination-the-use-of-artificial-intelligence-by-right-wing-extremists*

Gilbert, D. (2024, February 21). *Gab's racist AI chatbots have been instructed to deny the Holocaust*. Wired. *https://www.wired.com/story/gab-ai-chatbot-racist-holocaust/*

Hall, J. (2025, July). *The Terrorism Acts in 2023* [Corporate report]. Independent Reviewer of Terrorism Legislation (UK Government). *https://www.gov.uk/government/publications/the-terrorism-acts-in-2023*

Kajjo, S. (2024, May 23). *IS turns to artificial intelligence for advanced propaganda amid territorial defeats.* VOA News. *https://www.voanews.com/a/is-turns-to-artificial-intelligence-for-advanced-propaganda-amid-territorial-defeats/7624397.html*

Makuch, B. (2025, July 8). *How terrorist groups are leveraging AI to recruit and finance their operations.* The Guardian. *https://www.theguardian.com/world/2025/jul/08/terrorist-groups-artificial-intelligence*

The Middle East Media Research Institute (MEMRI). (2022, December 7). *ISIS supporter: AI lists steps for establishing caliphate*. *https://www.memri.org/cjlab/artificial-intelligence-lists-steps-establishing-caliphate-claims-isis-supporter*

Molas, B., & Lopes, H. (2024, October). *"Say it's only fictional": How the far-right is jailbreaking AI and what can be done about it* (Long read). International Centre for Counter-Terrorism. *https://icct.nl/sites/default/files/2024-10/Molas%20and%20Lopes.pdf*

Nelu, C. (2024, June 10). *Exploitation of generative AI by terrorist groups* [Short read]. The International Centre for Counter-Terrorism (ICCT). *https://icct.nl/publication/exploitation-generative-ai-terrorist-groups*

Piper, K. (2025, July 11). *Grok's MechaHitler disaster is a preview of AI disasters to come*. Vox. *https://www.vox.com/future-perfect/419631/grok-hitler-mechahitler-musk-ai-nazi*

Quinn, B., & Milmo, D. (2024, August 2). *How TikTok bots and AI have powered a resurgence in UK far-right violence.* The Guardian. *https://www.theguardian.com/politics/article/2024/aug/02/how-tiktok-bots-and-ai-have-powered-a-resurgence-in-uk-far-right-violence*

Stalinsky, S., Purdue, S. A., Sosnow, R., Smith, A., Agron, A., Szerman, N., Rosenfeld, N., Sloane, H., Strandberg, A., Hughes, J., Lee, K., & Avraham, L. (2024, June 20). *Neo-Nazis and white supremacists globally look to artificial intelligence to promote their message, spread misinformation, and aid their cause, January 2023–May 2024*. MEMRI. *https://www.memri.org/dttm/neo-nazis-and-white-supremacists-globally-look-artificial-intelligence-promote-their-message-0*

Tech Against Terrorism. (2023, November). *Early terrorist experimentation with generative artificial intelligence services* [Briefing]. *https://techagainstterrorism.org/hubfs/Tech%20Against%20Terrorism%20Briefing%20-%20Early%20terrorist%20experimentation%20with%20generative%20artificial%20intelligence%20services.pdf*

United Nations Office of Counter-Terrorism (UNOCT), & United Nations Interregional Crime and Justice Research Institute (UNICRI). (2021). *Algorithms and terrorism: The malicious use of artificial intelligence for terrorist purposes* [Report]. *https://unicri.org/News/Algorithms-Terrorism-Malicious-Use-Artificial-Intelligence-Terrorist-Purposes*

Verma, P. (2024, May 17). *These ISIS news anchors are AI fakes*. Their propaganda is real. The Washington Post. *https://www.washingtonpost.com/technology/2024/05/17/ai-isis-propaganda/*

CENTRE FOR RESEARCH AND EVIDENCE ON SECURITY THREATS

CREST Security Review provides a gateway to the very best knowledge and expertise. Its articles translate academic jargon to 'so what' answers and illustrate how behavioural and social science can be used effectively in everyday scenarios.

THE CENTRE FOR RESEARCH AND EVIDENCE ON SECURITY THREATS

CSR is produced by the Centre for Research and Evidence on Security Threats (CREST). CREST is funded by the UK's Home Office to identify and produce social science that enhances their understanding of security threats and capacity to counter them. Its funding is administered by the Economic and Social Research Council (ESRC Award ES/V002775/1).

The design of the research, data analysis, and interpretation are completed independently by the authors and should not be taken as representative of views held by those who fund CREST.

CREST has established a growing international network of over 200 researchers, commissioned research in priority areas, and begun to tackle some of the field's most pressing questions.

> '*CREST Security Review is a fantastic means by which we can keep practitioners, policy-makers and other stakeholders up-to-date on the impressive social and behavioural science occurring not only at CREST, but around the world.*'
>
> Professor Stacey Conchie, CREST Director

For more information on CREST and its work visit **www.crestresearch.ac.uk**

©2025 CREST Security Review | *CSR* is available under a Creative Commons 4.0 BY-NC-SA licence.
For more information on how you can use our content, visit: www.crestresearch.ac.uk/copyright