Counter-Terrorism

The Ethical Issues

Miller, S.R.M.; Henschke, A.H.; Feltes, J.

DOI
10.4337/9781800373075

Publication date
2021

Document Version
Final published version

Citation (APA)
Miller, S. R. M., Henschke, A. H., & Feltes, J. (Eds.) (2021). *Counter-Terrorism: The Ethical Issues*. Edward Elgar Publishing. https://doi.org/10.4337/9781800373075

Counter-Terrorism

Counter-Terrorism

The Ethical Issues

Edited by

Seumas Miller

Professor of Philosophy, Charles Sturt University, Australia, Delft University of Technology, the Netherlands and the Oxford Uehiro Centre for Practical Ethics, University of Oxford, UK

Adam Henschke

Senior Lecturer, Crawford School of Public Policy, Australian National University, Australia

Jonas Feltes

PhD candidate, Department of Values, Technology & Innovation, Delft University of Technology, the Netherlands

Cheltenham, UK • Northampton, MA, USA

Published by
Edward Elgar Publishing Limited
The Lypiatts
15 Lansdown Road
Cheltenham
Glos GL50 2JA
UK

Edward Elgar Publishing, Inc.
William Pratt House
9 Dewey Court
Northampton
Massachusetts 01060
USA

A catalogue record for this book
is available from the British Library

Library of Congress Control Number: 2021938668

This book is available electronically in the Elgaronline
Political Science and Public Policy subject collection
http://dx.doi.org/10.4337/9781800373075

ISBN 978 1 80037 306 8 (cased)
ISBN 978 1 80037 307 5 (eBook)

Contents

Contributors

Mary B. DeRosa is Professor from Practice at Georgetown Law, where she focuses on national security law and practice, and she is Co-Director of the Global Legal Scholars Program. Previously, she served as Deputy Counsel to the President and National Security Council (NSC) Legal Adviser in the Obama Administration and as NSC Legal Adviser in the Clinton Administration. She also served as Chief Counsel for National Security on the Senate Judiciary Committee, Special Counsel at the Department of Defense, and was a member of the President's Intelligence Advisory Board and legal advisory boards at the National Security Agency and the Central Intelligence Agency.

Jonas Feltes is a PhD candidate in the European Research Council's Advanced Grant, 'Global Terrorism and Collective Moral Responsibility: Redesigning Military, Police and Intelligence Institutions in Liberal Democracies' at the Delft University of Technology. He has an academic background in both the history of terrorism and the ethics and history of technology. Jonas studied history, philosophy and the history and philosophy of technology in Germany and the Netherlands and specialised in the history and ethics of technology and (global) security. In previous work, he investigated the relation between selected explosive technologies and terrorism.

Dr Adam Henschke is a senior lecturer with the Crawford School of Public Policy at the Australian National University in Canberra, Australia. He is an applied ethicist and works on areas where technology, ethics and national security policy intersect. His research concerns ethical and philosophical analyses of information technology and its uses, military ethics and relations between ethics and national security. He has published on surveillance, emerging military technologies and intelligence and cyberspace. He is also interested in moral psychology, experimental philosophy and their relations to decision-making and policy development.

Professor Seumas Miller holds research positions at Charles Sturt University, Canberra, Delft University of Technology and the University of Oxford. He is the author or co-author of 20 books, including *Terrorism and Counter-Terrorism: Ethics and Liberal Democracy* (Blackwell, 2009), *Shooting to Kill: The Ethics of Police and Military Use of Lethal Force* (Oxford University Press, 2016) and *Institutional Corruption* (Cambridge

University Press, 2017), and over 200 academic articles. He is the Principal Investigator on the European Research Council's Advanced Grant, 'Global Terrorism and Collective Moral Responsibility: Redesigning Military, Police and Intelligence Institutions in Liberal Democracies' (GTCMR. No. 670172, http://www.counterterrorismethics.com).

Professor Mitt Regan is McDevitt Professor of Jurisprudence and Co-Director of the Center on National Security and the Law at Georgetown University Law Center, and Senior Fellow at the Stockdale Center on Ethical Leadership at the United States Naval Academy. His work focuses on legal and philosophical issues relating to international law, national security, counter-terrorism and human rights.

Dr Scott Robbins is a postdoctoral researcher in the ethics of artificial intelligence at the Center for Advanced Security, Strategic and Integration Studies at Bonn University in Germany. He recently completed his PhD in the ethics of machine learning in a counter-terrorism context at the Delft University of Technology. Scott has a BSc in Computer Science from California State University, Chico, and an MSc in Ethics of Technology from the University of Twente. He is a founding member of the Foundation for Responsible Robotics and a member of the 4TU Centre for Ethics and Technology. Scott is sceptical of AI as a grand solution to societal problems and argues that AI should be boring.

Dr Michael Robillard is a postdoctoral research fellow with the European Research Council's Advanced Grant on 'Global Terrorism and Collective Moral Responsibility: Redesigning Military, Police and Intelligence Institutions in Liberal Democracies'. He received his PhD from the University of Connecticut in 2016. Prior to that, he was a resident research fellow at the Stockdale Center for Ethical Leadership at the United States Naval Academy. Michael's past research has focused on the overlap between normative theories of exploitation and present-day military recruitment. He has also written on the ethics of autonomous weapons, as well as war and its relation to future generations. Michael is an Iraq War veteran, United States Military Academy graduate and former Airborne Ranger.

Dr Michael Skerker is Associate Professor in the Leadership, Ethics, and Law Department at the United States Naval Academy. His academic interests include professional ethics, particularly, police, military and intelligence ethics. His publications include *An Ethics of Interrogation* (University of Chicago Press, 2010), *The Moral Status of Combatants* (Routledge, 2020) and *Military Virtues* (Howgate, 2019). More information can be found on his web-page (https://www.usna.edu/LEAD/DivisionStaff/Skerker_Michael_NE203.php).

Levi J. West is Director of Terrorism Studies at the Australian Graduate School of Policing and Security at Charles Sturt University, Canberra, and a PhD scholar at Victoria University, Melbourne. His research interests include the role of technology in terrorism, radicalisation and terrorism strategies and tactics.

Alexandra L. White is an Associate at Sidley Austin in Washington, DC. She received her JD, *cum laude*, from Georgetown University Law Center in 2020, where she was also a Global Law Scholar. She holds an AB from Brown University in International Relations with a focus on security and conflict studies.

Acknowledgement

This research was conducted under the auspices of the European Research Council's Advanced Grant program as part of the grant titled, 'Global Terrorism and Collective Moral Responsibility: Redesigning Military, Police and Intelligence Institutions in Liberal Democracies' (GTCMR. No. 670172); Principal Investigator, Professor Seumas Miller.

We wish to thank the editors of the following academic publications of Seumas Miller for use of some of the material contained therein: 'Collective Moral Responsibility: An Individualist Account' *Midwest Studies in Philosophy*, vol. XXX. Boston: Wiley-Blackwell, 2006; *Terrorism and Counter-Terrorism: Ethics and Liberal Democracy*. Oxford: Blackwell, 2009; 'The Moral Justification for Preventive Detention' *Criminal Justice Ethics* 37: 2, 2018.

Abbreviations

AI	Artificial Intelligence
AN	Ammonium nitrate
ANC	African National Congress
AQ	al-Qaeda
AUMF	Authorization for Use of Military Force against Terrorists
BDA	Battle damage assessment
BDC	Bulk data collection
CBRN	Chemical, biological, radiological and nuclear
CDE	Collateral damage estimate
ChemVerbotsV	Chemikalien-Verbotsverordnung
CIVIC	Center for Civilians in Conflict
CNA	Center for Naval Analyses
CRIN	Context-relative informational norms
CT	Counter-terrorism
DNI	Director of National Intelligence
DoD	Department of Defense
ECHR	European Convention on Human Rights
FATA	Federally Administered Tribal Areas
FBI	Federal Bureau of Investigation
FISA	Foreign Intelligence Surveillance Act
FRT	Facial recognition technology
FTO	Foreign Terrorist Organization
GCHQ	Government Communications Headquarters
GCS	Ground control station
GefStoffV	Gefahrstoffverordnung
GIFCT	Global Internet Forum for Countering Terrorism
GTAZ	*Gemeinsames Terrorismusabwehrzentrum*

HVT	High-value target
IC	Intelligence community
ICCPR	International Covenant on Civil and Political Rights
ICRC	International Committee of the Red Cross
ICT	Information and communications technology
IHL	International humanitarian law
IRA	Irish Republican Army
ISAF	International Security Assistance Force
ISIS	Islamic State of Iraq and Syria
JIT	Just Intelligence Theory
JWT	Just War Theory
LOAC	Laws of armed conflict
LRE	Launch and recovery element
MEC	Moral equality of combatants
NMH	No Means to Harm
NSA	National Security Agency
NSC	National Security Council
PID	Positive identification
POTUS	President of the United States
POW	Prisoner of war
PPE	Personal protective equipment
PPG	Presidential Policy Guidance
PSYOP	Psychological operations
SOCMINT	Social media intelligence
SUE	Strategic Use of Evidence
UK	United Kingdom
UN	United Nations
US	United States

Introduction to counter-terrorism: the ethical issues

Unquestionably, terrorism has emerged as one of the most important and pressing problems confronting contemporary societies and, in particular, liberal democracies. However, it has also emerged as a site of ideological contestation. In this context, it is crucial that clear-headed, objective and academically rigorous analyses consider: (1) the defining characteristics of terrorism, and of salient terrorist strategies and tactics, for example, the deliberate killing of innocent civilians and the use of social media to recruit members and spread terrorist ideology; and (2) the moral justifications (or lack thereof) of the various counter-terrorism measures taken, for example, targeted killing, enhanced interrogation, preventive detention and bulk metadata collection.

The volume consists of analyses of central ethical issues that have arisen in combating global terrorism and, in particular, jihadist terrorist groups, notably al-Qaeda, the Islamic State and their affiliates. A number of the chapters deal with some of the theoretical issues that arise in relation to terrorism, such as the definition of terrorism and the concept of collective responsibility. However, most of the chapters are concerned with specific ethical issues. Hence, topics include targeted killing, enhanced interrogation of terrorists, preventive detention, freedom of expression and terrorist content on social media, psychological warfare, bulk metadata collection and responses to terrorist attacks that use weapons of mass destruction. The analyses are anchored in 'reality' by using specific examples of terrorist organisations, tactics and ends. Here follows a brief description of each of the chapters in this volume.

1. MITT REGAN AND ALEXANDRA L. WHITE, 'PREVENTIVE CRIMINAL LAW: TERRORIST CRIMES AND LIBERAL DEMOCRATIC VALUES'

Spurred by United Nations Security Council resolutions and directives of the European Parliament, liberal democracies since 9/11 have enacted several criminal statutes that establish new offences related to terrorism, or that increase penalties for the commission of already outlawed crimes when such crimes are committed in relation to terrorism. This trend reflects the belief

that terrorism represents a threat that is more significant than the threat from ordinary criminal activity, and that must be deterred by distinctive means at an earlier stage in the process or through the threat of heightened sentencing. It also reflects a growing concern about how individuals are being recruited and radicalised today, which oftentimes includes leaving one's home country to travel to join a terrorist group for training or to engage directly in conflict. The distinctive measures that states have taken to address these concerns include criminalising conduct at an earlier stage in the commission of a crime than is the case with other criminal offences; characterising various types of behaviour as preparation for or assistance to terrorist acts; imposing severe penalties for crimes that are designated as constituting terrorism; and using terrorism as the basis for engaging in expanded surveillance and investigatory activities, restricting immigration and imposing preventive detention and other restrictive measures. As these measures have proliferated, some observers have raised concerns that they risk inconsistency with liberal democratic values. The absence of a common definition of terrorism, for instance, could lead to definitions that suppress legitimate dissent, and result in substantially different treatment across states for the same behaviour. Criminalisation of behaviour deemed to be preparatory could unjustifiably infringe on rights of speech, opinion and association in some cases, a risk that can be especially high if culpable behaviour is defined in vague and open-ended ways that leave considerable discretion to law enforcement or judicial officials. Broad definitions also pose potential problems when putative terrorist-related activities are the basis for expanded surveillance, investigative powers and more severe penalties than for ordinary crime. This chapter describes the various terrorist provisions that have been adopted by Australia, Belgium, France, Germany, the Netherlands, the United Kingdom and the United States; the information available about how they have been applied; and the extent to which they have the potential to raise concerns about intrusion on individual rights that are deemed important in liberal democracies. It then suggests how states might approach the task of enacting terrorism-related offences so as to minimise the risks of such intrusion.

2. SEUMAS MILLER AND JONAS FELTES, 'THE DEFINITION OF TERRORISM'

It is often suggested that 'one man's terrorist is another man's freedom fighter', but defining terrorism is, in fact, both possible and desirable, for only then can the term 'terrorist' cease to be used purely in the context of ideological name-calling. A number of academically serious definitions of terrorism are already on offer. These definitions tend to fall into two camps. Some define terrorism, in part, in terms of killing innocent persons. Others define terrorism,

in part, in terms of killing civilians, or at least some category of persons that is not by definition or not necessarily innocent. A feature of most definitions of terrorism, irrespective of which camp they belong to, is the failure to specify which of the necessary conditions that constitute the definition, including the political effects, have to be intended and realised for the action to count as terrorism. For instance, the intentional killing of an innocent person in the service of a political purpose would normally count as an act of terrorism, but what if the action goes unreported, is intended to go unreported and, therefore, fails to have any public political impact, except sending its intended message to members of the security forces? In this chapter, a definition of terrorism will not only be elaborated and defended against rival accounts, it will also consider questions of detail that are neglected in most definitions. Specifically, it is argued that terrorism is a strategy that: (1) consists of state or non-state actors deliberately performing acts of violence aimed at (directly or indirectly) seriously harming persons who are not military combatants, human rights violators or violent revolutionaries; (2) consists of violent actions that ought to be criminalised; (3) is an intended means of terrorising the members of some social, economic, ethnic, political, or other group to achieve a political purpose; and (4) relies on the violence receiving a degree of publicity, at least to the extent necessary to engender widespread fear in the target group.

3. SEUMAS MILLER AND JONAS FELTES, 'COLLECTIVE RESPONSIBILITY AND COUNTER-TERRORISM'

The chapter begins with an account of collective moral responsibility elaborated and defended elsewhere, namely, collective responsibility as joint responsibility (Seumas Miller, 'Collective Responsibility: An Individualist Account', *Midwest Studies in Philosophy*, 2006). The chapter then proceeds with an application of this notion of collective responsibility to the cooperative, that is, joint actions of terrorists, and also to the joint actions (or joint omissions) of members of security agencies that are, inter alia, engaged in counter-terrorism. This theoretical notion of collective moral responsibility is also used in later chapters, for example, on counter-terrorism and weapons of mass destruction. Collective responsibility of the kind in question here is the responsibility that attaches to the participants of a joint action for the performance of that joint action and, in particular, for the realisation of the collective end of the joint action. There are different accounts of collective responsibility, some of which pertain to the responsibility of groups and organisations per se for their group or 'corporate' (so to speak) actions. Here, our concern is only with collective responsibility for joint actions of human beings in their capacity as institutional role occupants. On the view of collective responsibility as joint responsibility,

collective responsibility is ascribed to individual human beings only, albeit jointly. Moreover, institutional actors can be ascribed collective institutional responsibility when they act jointly in accordance with their institutional roles.

An effective counter-terrorism strategy, it is argued, involves cooperation between multiple security and other state agencies, financial institutions and other businesses, and members of the public. Indeed, it includes what has in other contexts been termed a 'web of prevention' (see also Chapter 14). As such, it comprises multiple, coordinated, layered structures of joint action. Moreover, since the web of prevention also has a diachronic dimension that consists of the operation of institutional processes in which multiple institutional actors function in accordance with a division of labour, it involves complex, intersecting chains of responsibility.

4. MICHAEL ROBILLARD, 'KILL, WOUND OR CAPTURE: ETHICS CONSIDERATIONS FOR COUNTER-TERRORISM OPERATIONS'

In this chapter, some of the major moral and pragmatic elements of the kill, wound or capture criteria for counter-terrorism operations are identified and analysed. Just War Theory, broadly construed, and particularly facets of *jus ad bellum* and *jus in bello*, are explored in relation to counter-terrorism in particular. Furthermore, existing legal guidance on targeted killing for the United States and internationally is investigated. Lastly, a set of other under-acknowledged normative factors pertinent to kill, wound or capture operations are discussed. While these moral considerations are not intended to be exhaustive when it comes to decision-making for kill, wound or capture operations, nonetheless, they are offered to augment existing counter-terrorism thinking and planning because they are not often explicitly expressed in present targeted killing guidelines. It is, therefore, left up to present and future counter-terrorism commanders and operators to decide what level of stringency makes the most sense to them given other competing ethical and strategic priorities.

5. MARY B. DEROSA AND MITT REGAN, 'ACCOUNTABILITY FOR TARGETED KILLING'

The use by the United States of remotely piloted aircraft to engage in targeted killing outside of active combat has generated intense controversy for more than a decade. In response to criticism, the United States in recent years has restricted these operations as a matter of policy, although not law, in accordance with standards that approximate human rights principles. These require that individuals pose a threat to the United States, that capture of them

is infeasible, and that there is minimal risk of innocent civilian casualties. These requirements arguably bring the program more closely in line with the demands of ordinary morality with regard to state use of lethal force. Critics maintain, however, that there is no way to hold the United States accountable under these principles because there is no disclosure of how they are applied in particular targeted strikes. This failure of accountability means that the program does not satisfy a fundamental ethical condition on a state's use of lethal force. This chapter analyses concerns about accountability in the United States targeted killing program with respect to: (1) the determination of which persons are put on a targeting list, and (2) how individual missions are carried out. It first provides a detailed description of the decision-making process, and the mechanisms designed to ensure accountability, at each of these stages. It then assesses the effectiveness of these mechanisms and evaluates other arrangements that proponents claim will provide more meaningful accountability. The analysis in this chapter thus illuminates the extent to which it is possible for targeted killing programs to be conducted in accordance with the moral demands of accountability.

6. MICHAEL SKERKER, 'INTERROGATION ETHICS IN COUNTER-TERROR OPERATIONS'

Recent research has led to an emerging scientific consensus about best practices in interrogation. Government agencies in Norway, the United States, the United Kingdom and other commonwealth countries have begun to train personnel in scientifically validated, rapport-based interrogation methods that are practical and moral improvements on older methods that seek to overcome or circumvent the interrogatee's will through emotional pressure or trickery. This chapter presents four modern types of interrogation and assesses them from practical and moral perspectives in a counter-terror context. The approach in this chapter takes into account suspected terrorists' rights, the protective duties of interrogators and concerns for the psychological and moral health of the interrogators themselves.

7. SEUMAS MILLER, 'PREVENTIVE DETENTION OF TERRORISTS'

One of the most problematic forms of detention from an ethical perspective is preventive detention. Preventive detention is a coercive measure that is, under normal circumstances, a violation of individual freedom. However, some have argued that it may be justified as a counter-terrorist measure. Arguably: (1) terrorists are, by definition, guilty of crimes since they murder innocent civilians and, therefore, should be subjected to criminal procedures

and punished accordingly; and (2) terrorists who are members of organisations engaged in protracted armed conflict may well be de facto combatants (terrorist-combatants) and, as such, if captured can reasonably be incarcerated until the cessation of hostilities; the same point holds for members of non-terrorist insurrectionary groups. This chapter discusses the arguments for and against the preventive detention of terrorists.

8. SEUMAS MILLER, 'USE OF STINGS IN COUNTER-TERRORISM: ENTRAPMENT AND ETHICS'

While we need to distinguish the conditions that define entrapment defences from those under which counter-terrorism stings might be morally – and ought to be legally – justified, the former are a subset of the latter. Thus, the so-called subjective and objective tests used in the United States in relation to the legal defence of entrapment provide a useful initial guide to the discussion of the wider ethical issues raised by stings, including counter-terrorism stings and, in particular, the issues of 'creating crime' and (relatedly) of injustice. Presumably, the target of a successful sting who is convicted of terrorism has been unjustly treated if they did not commit, and would not have committed, an act of terrorism, absent the sting. After all, in these circumstances, the only crime (if crime it is) that has been, or will be, committed is the one manufactured by the sting. Of course, in addition to the problem of the injustice to the target, there is the matter of prevention. The primary purpose of stings is to prevent crime and, in the cases of interest to us here, prevent terrorist attacks. But if the target of a sting did not and would not have committed an act of terrorism (absent the sting), then obviously the primary purpose of the sting has not been achieved, since no terrorist attack has been prevented (other than, perhaps, the one manufactured by the sting operation).

9. LEVI J. WEST, 'COUNTER-TERRORISM, SOCIAL MEDIA AND THE REGULATION OF EXTREMIST CONTENT'

This chapter analyses the ethical issues raised by the policies and practices that have developed, and continue to evolve, as a result of the emergence of the use of various forms of social media and contemporary information and communications technology (ICT) by terrorist entities. The emergence of the Islamic State has evidenced the power and effectiveness of the adroit exploitation of social media platforms as a vector for terrorist propaganda and radicalisation, and for remote command and control of operations. As a result, counter-terrorism policies and practices have evolved to respond to this chal-

lenge. In assessing these responses, this chapter provides a brief overview of the mechanisms by which terrorist entities have come to exploit social media and ICT, before providing an analysis of a number of the ethical dilemmas raised by the counter measures. The chapter analyses the appropriateness of the emergence of the censoring of content by private, for-profit corporations, rather than by the nation state; it also explores the broadening collection and surveillance role of social media companies and the disproportionate responses to content based on ideological characteristics. In doing so, this chapter identifies how the responses to the challenge of the use of social media and ICT by terrorist entities warrants substantial consideration from an ethical perspective, given the risks posed by the existing policies and practices, and the need to ensure that future policies and practices are informed by ethical considerations.

10. ADAM HENSCHKE, 'ON FREE PUBLIC COMMUNICATION AND TERRORISM ONLINE'

The issue of freedom of speech and counter-radicalisation is evident from, among other things, the calls to delete all terrorist propaganda from websites and regulate social media. Liberals (following, for instance, J.S. Mill in his classic study, *On Liberty*) tend to hold it to be axiomatic that the right to freedom of speech and thought is inconsistent with laws prohibiting sedition and, more generally, propagating political ideologies. Specifically, it is assumed that a fundamental feature of any well-ordered liberal democracy is that its citizens have a right to argue for, and disseminate, the view that the political system and/or the government of the day ought to be overthrown by peaceful or, if necessary, by violent means. This is consistent with their being, for example, laws against inciting unruly mobs to violence against politicians or police; disseminating information that would enable others to overthrow the government, for example, how to construct and set off a nuclear device; or a person in authority ordering subordinates to engage in violent action against the state, for example, a military or police officer directing subordinates to engage in acts of terrorism. Consistent with this last point, it might be that the nature of the authority relationship between some fundamentalist Muslim leaders and their followers is such that the latter are subordinates in an appropriate sense – that is, they will, if directed, engage in terrorist acts. If so, then laws against the issuing of 'directives' – for example, fatwahs – by Muslim leaders might not constitute an infringement of the right to free speech. This chapter discusses freedom of expression in countering terrorist propaganda and the like on social media.

11. MICHAEL ROBILLARD, 'COUNTER-TERRORISM AND PSYOP'

The ethics of war is often delineated in terms of *jus ad bellum* (the ethics of going to war) and *jus in bello* (ethical behaviour in war). With the advent of the informational age, the increased use of informational warfare and PSYOP (psychological operations) has significantly problematised standard *ad bellum* as well as *in bello* thinking. Without a physical territorial boundary being crossed, without a designated three-dimensional battle-space, and without physical infrastructure or human bodies taking noticeable kinetic damage, it is unclear how PSYOP fits within contemporary just war thinking. This is made more problematic when it comes to the ethics of counter-terrorism. This chapter gives an explanation of contemporary PSYOP practices related to counter-terrorism and articulates various values, tensions and trade-offs connected to these practices. Borrowing from LTC Bob Underwood, the notion that non-kinetic communicative effects are metaphysically and ethically inseparable from kinetic actions is advanced. In other words, PSYOP will often cause foreseeable and unforeseeable second-order kinetic effects. Conversely, it is argued that kinetic operations often generate foreseeable and unforeseeable second-order communicative effects. From these dual notions, ethical and efficacious counter-terrorism operations ought to regard these communicative effects as one of the highest strategic priorities.

12. ADAM HENSCHKE, 'FROM "NEED TO SHARE" TO "NEED TO CARE": INFORMATION AGGREGATION AND THE NEED TO CARE ABOUT HOW SURVEILLANCE TECHNOLOGIES ARE USED FOR COUNTER-TERRORISM'

Technological innovation is disrupting how we use and treat information. These disruptions affect both theory and practice. The key point of this chapter is to argue that one of the foundational shifts brought about by convergent information technologies is the ease with which information can be aggregated. To explain this point, a scenario is developed in which what seems like an innocuous set of activities ends up having significant implications for security. It is then shown that information aggregation is one of the key steps. This focus on aggregation reveals how theory – in this case, privacy – and practice – in this case, intelligence – are both affected. Some suggestions are made on how disaggregating information can help resolve issues in privacy, intelligence and the tensions that sometimes arise between civil liberties and national security.

13. SCOTT ROBBINS, 'BULK DATA COLLECTION, NATIONAL SECURITY AND ETHICS'

The ethics of intelligence has frequently been discussed by contemporary scholars within a Just War Theory framework – those principles deemed necessary for the ethical initiation, conduct and termination of war. Principles such as just cause, right intention, proportionality, last resort and others are now being used to ethically evaluate intelligence practices. These scholars are aware that war and intelligence practices are not the same kind of activity (for example, war is kinetic) and have made efforts to modify Just War Theory into a Just Intelligence Theory that accounts for the differences. The focus in this chapter is the application of some of the latest work in Just Intelligence Theory to bulk data collection. This serves two purposes: first, to come to an understanding of the important ethical issues surrounding the practice of bulk data collection for intelligence purposes and, second, to highlight how Just Intelligence principles can be used to evaluate a specific intelligence program.

14. JONAS FELTES, 'COLLECTIVE MORAL RESPONSIBILITY AND CHEMICAL, BIOLOGICAL, RADIOLOGICAL AND NUCLEAR TERRORISM: THE CASE OF PHOSPHINE'

Chemical weapons have been used by terrorist state actors, such as Saddam Hussein and Bashar al-Assad, in Syria, Iraq and elsewhere. However, biological and nuclear weapons are also potential, if as-yet-unused, forms of weaponry that can be accessed by non-state-actor terrorist groups. Evidently, there is a non-negligible threat of the use of chemical, biological, radiological and nuclear (CBRN) agents by terrorist groups, and it is likely to increase rather than decrease. This chapter is concerned with the counter-terrorism response to CBRN terrorism and, in particular, the need to establish and maintain a so-called web of prevention. Such a web of prevention involves the taking of a variety of integrated counter-terrorism measures not only by security agencies but also by commercial firms and individual citizens, for example, reporting requirements, hence, the relevance of the notion of collective moral responsibility elaborated in Chapter 3 above. Such a web of prevention is, in essence, an institutionalisation of the collective responsibility to avert or mitigate the catastrophic effects of a CBRN attack. This chapter will illustrate the need for such a web of prevention by discussing the toxic gas phosphine.

1. Preventive criminal law: terrorist crimes and liberal democratic values

Mitt Regan and Alexandra L. White

1. INTRODUCTION

While state concern about terrorism is not new, it has intensified since the attacks on the United States on 11 September 2001. Spurred by the United Nations (UN), one response has been state creation of criminal offences specifically related to terrorism. This reflects an effort both to single out terrorism as an especially egregious crime, as well as to interdict it at an earlier stage than is the case for other crimes. These measures typically focus on behaviour that ordinarily is not criminal in itself, but that is seen as preparatory for, or in support of, eventual terrorist attacks. In this respect, their aim is preventive in that they are 'directed toward crimes as yet uncommitted' (Jarvis and Legrand 2018, p. 200).

This chapter examines the extent to which such measures may be in tension with liberal democratic ethical values, and suggests how states can resolve this tension. To illuminate these issues, we discuss terrorist offences adopted by Australia, Belgium, France, Germany, the United Kingdom, and the United States. Our analysis focuses on the potential implications of these offences for the rights of speech, travel, association, and due process, and suggests how laws may be framed so that they are consistent with the liberal democratic values that are the foundation of these rights.

The increase in human rights conventions in the period after World War II means that ethical values and legal principles are often closely aligned in liberal democracies in the form of individual rights. The analysis in this chapter of the implications of preventive criminal law for liberal democratic ethical values therefore focuses mainly on the extent to which criminal measures risk violating the human rights that reflect these values. Where appropriate, however, we will note when legal compliance may not fully honour fundamental ethical principles in liberal democracies.

2. INTERNATIONAL INITIATIVES

Shortly after 9/11, UN Security Council Resolution 1373 directed all member states to ensure that financing, planning, preparation, or perpetration of terrorist acts and supporting terrorist acts 'are established as serious criminal offences in domestic laws and regulations and that the punishment duly reflects the seriousness of such terrorist acts' (UN 2011). In 2014, Resolution 2178 reflected concern about persons returning to their native countries after fighting on behalf of terrorist organisations. It mandated that all states enact laws prohibiting: (1) certain travel 'for the purpose of perpetration, planning, or preparation of, or participation in terrorist acts, or the providing or receiving of terrorist training'; (2) terrorism financing; and (3) the organisation of travel for, or recruitment for travel abroad of, prospective foreign fighters (UN 2014, pp. 4–5). Human Rights Watch estimated that in the two years following Resolution 2178, more than forty-five states enacted counter-terrorism measures addressing foreign fighters (Tayler 2016, p. 9).

In 2017, the European Parliament issued a directive on combating terrorism that instructed member states of the European Union to criminalise several activities related to terrorism (European Parliament 2017). These included: directing a terrorist group; participating in the activities of a terrorist group by knowingly providing assistance to it; public provocation to commit a terrorist offence; recruitment for terrorism; providing or receiving training for terrorism; travelling for the purpose of terrorism; organising or otherwise facilitating travel for the purpose of terrorism; aiding, abetting, inciting, and attempting terrorism; and committing theft or extortion, or falsifying documents in connection with terrorist activities (European Parliament 2017). Most states have followed these directives to establish criminal offences related to terrorism. Some of these offences criminalise preparatory conduct at an earlier point prior to commission of a crime than occurs for non-terrorism-related crimes, while others add new crimes related to terrorism, and still others significantly increase punishment for existing crimes if they are committed in connection with terrorism. States that adopt such measures remain obligated to respect and ensure the enjoyment of human rights under instruments such as the International Covenant on Civil and Political Rights (ICCPR) and the European Convention on Human Rights (ECHR). The remainder of the chapter discusses how the creation of terrorist crimes may raise ethical issues by virtue of their incompatibility with the liberties codified in these instruments.

3. VAGUENESS AND FORESEEABILITY

Criminal laws that do not contain, or are not accompanied by, a clear definition of terms, such as 'terrorist' or 'terrorist intent', as well as laws that do not clearly define what acts constitute terrorism, may be impermissibly vague and in violation of the principle that individuals must be able to reasonably foresee that their conduct will be criminal before they are subject to prosecution. Both Article 15 of the ICCPR (UN 1976) and Article 7 of the ECHR (Council of Europe 1952) prohibit punishment for an action that was not an offence at the time it was committed, stating that 'no one shall be held guilty of any criminal offence on account of any act or omission which did not constitute a criminal offence under national or international law at the time when it was committed' (Council of Europe 1952). According to the European Court of Human Rights (2020, p. 12), the doctrine of foreseeability does not require absolute clarity in crafting criminal laws. Conduct is foreseeable in criminal law if the measure that prohibits it can be defined with sufficient clarity through interpretation by courts or lawyers (European Court of Human Rights 2020, p. 12).

One foreseeability challenge is that there is no international consensus on the meaning of terms, such as 'terrorism' or 'terrorist'. As Ben Saul (2015) has noted, 'Excessively wide or loose concepts of terrorism can seriously jeopardise internationally protected human rights'. In Resolution 1566 (UN 2004), the UN Security Council described terrorism as:

> criminal acts, including against civilians, committed with the intent to cause death or serious bodily injury, or taking of hostages, with the purpose to provoke a state of terror in the general public or in a group of persons or particular persons, intimidate a population or compel a government or an international organisation to do or to abstain from doing any act.

States are not required, however, to adopt this definition. Indeed, the UN Office on Drugs and Crime has offered states five possible definitions of terrorist acts in its Model Legislative Provisions against Terrorism (Article 15 2009).

The 2017 European Directive offered a definition of 'terrorist offenses' (European Parliament 2017). Article 3 provides that certain crimes are to be regarded as terrorist offences if they are committed with the aim of:

(a) seriously intimidating a population;
(b) unduly compelling a government or an international organisation to perform or abstain from performing any act; or
(c) seriously destabilising or destroying the fundamental political, constitutional, economic, or social structures of a country or an international organisation (European Parliament 2017).

Despite such regional efforts to establish a standard definition, '[c]oncepts of terrorism in national law are…startlingly diverse and there is little evidence of global convergence, and certainly as yet no customary international crime of terrorism' (Saul 2015, p. 2). In its review of state implementation of Resolution 1373, the UN noted that 'national laws of a number of States criminalise terrorist acts in vague or overbroad terms that could lead to abuse' (UN 2016, p. 124).

Vagueness can also be a concern regarding preparatory offences. As Robert Chesney (2007) has noted, 'The farther that one moves from the paradigm of a completed act—as one moves backwards successively through attempt, to advanced planning, to initial planning, and so forth—the more tenuous the link between the defendant and the anticipated harm becomes' (p. 435).

The United Kingdom's Terrorism Act of 2000 makes it a crime if a person, in relation to an act of terrorism, 'engages in any conduct in preparation for giving effect to his intention' (UK Chapter 11, Section 5). Similarly, Australia's criminal code makes unlawful 'any act in preparation for, or planning, a terrorist act' (Australia Criminal Code 101.6), and prohibits possessing a thing connected with preparing or assisting a terrorist act (101.4), collecting or making a document connected to preparing or assisting with a terrorist act (101.5), and preparing to travel 'to commit hostile acts' (119.4).

Other types of preparatory acts that were criminalised include collecting information about people or a place in order to carry out a terrorist act against the people or the place, an offence that both Belgium and France have similarly codified (Belgium code penal 140septies; French code penal 421-2-6). Related to gathering information, the United Kingdom has criminalised actions even further along the spectrum of preparatory action by making it unlawful to view certain terrorist material. Thus, it is an offence if a person views or accesses online, possesses, or collects information that is 'likely to be useful to a person committing or preparing an act of terrorism' if that person does not have a defence of, at the time of such action, not having had reason to believe that the material was going to contain information useful to committing a terrorist act, or if the person was a journalist or doing academic research (UK 58(1)). Such statutes, on their face, often provide minimal guidance on behaviour that would violate them. The unlawfulness of the conduct hinges, rather, on its connection to terrorist intention, which may be defined only in broad terms.

One effort to avoid challenges with foreseeability is to designate certain groups as terrorist organisations and to prohibit any form of assistance to them. The publication of groups that have been designated as such organisations then is presumed to provide clear notice that providing any support to them is a crime. As Jarvis and Legrand (2018, p. 200) note, this approach is widely used by states as well as several international governmental organisations, such as the UN and the European Union. In the United States, for instance, it is

sufficient for criminal liability that an individual simply knows that an organisation to which they provide material support has been designated a Foreign Terrorist Organisation (FTO) by the Secretary of State (18 USC §2339B). This is in contrast with another statute that makes it a crime to provide various forms of material support, 'knowing or intending that they are to be used in preparation for, or in carrying out' terrorist offences (18 USC §2339A).

4. FREEDOM OF EXPRESSION

Article 19 of the ICCPR protects the right to expression, although it says that the right may be subject to restrictions due to considerations of national security and public order (UN 1976). Similarly, Article 10 of the ECHR allows states to limit freedom of expression if it is for the purpose of national security in order to prevent 'disorder or crime' (Council of Europe 1952). At the same time, the UN emphasises the importance of narrow restrictions on speech. It has declared, 'Laws should only allow for the criminal prosecution of direct incitement to terrorism, that is, speech that directly encourages the commission of a crime, is intended to result in criminal action and creates a danger of criminal action' (UN 2018, p. 56).

Several states have enacted laws to criminalise certain types of speech because it could be considered preparatory, but the logic of such regulations falters where speech is criminalised even in the absence of an established connection to a terrorist act or terrorist intent. Many states, such as Australia (Anti-Terrorism Act of 2005), Belgium (Art. 140), and France (Art. 421-2-5), criminalise advocacy of terrorism but do so with varying levels of nexus between the speech and potential incitement to violence. Belgium does not require the speech to directly cause harm, but it does mandate that the disseminator of a message intends to incite the commission of a terrorist act (Art. 140). Australia broadly outlaws 'counselling, promoting, encouraging or urging others to commit a terrorism offence' (Art. 80.2C).

Judicial decisions in the United Kingdom have emphasised the impact of speech on terrorist recruitment and online radicalisation. For example, in 2015, a woman received a prison sentence of three and a half years for having tweeted 45 000 times in support of the Islamic State of Iraq and Syria (ISIS; *The Guardian* 2015). At trial, the judge told her, 'The material you were disseminating encouraged young men to go and fight and you now accept that was your intention' (*The Guardian* 2015). A year later, a young man was convicted for posting 8 000 messages on Twitter expressing support for ISIS (*The Guardian* 2016). The prosecution argued that the defendant engaged in 'a sustained effort indirectly to encourage others to engage in terrorism', and *The Guardian* (2016) reported that he 'portrayed terrorists as role models, referring to the "magnificent 19" hijackers from the September 11 suicide attacks'.

A more expansive law is Article 421-2-5 of France's penal code, which makes it a crime to publicly apologise for terrorism acts, which is often translated as 'glorifying' terrorism. This potentially limits speech based on its content without the intent or effect of inciting violence. The French Constitutional Council found in 2018 (Constitutional Council 2018) that the law does not infringe on freedom of expression. It found the law's intent was to prevent terrorism by prohibiting persons from 'broadcasting expressions endorsing acts that have the goal of seriously causing disturbance to the public order by intimidation or terrorism' (p. 20). It further found that public dissemination of such 'dangerous ideas and expressions' is sufficient in itself to disturb the public order, which permits its proscription (p. 21).

One notable conviction under the law has been of someone who posted a message on Facebook after a police officer was killed in a hostage exchange: 'Every time a police officer is killed, and it's not every day, I think of my friend Rémi Fraisse…This time it was a colonel—how great!' (Held 2018). Fraisse was killed by a police stun grenade while protesting a dam project in 2014 (Held 2018). Another conviction, resulting in a seven-month suspended sentence, was of a vegan activist who posted on social media after a butcher was killed in an attack on a supermarket: 'It shocks you that an assassin is killed by a terrorist? Not me, I have zero compassion for him. There is justice after all' (Houry 2018).

Although the comments that were the basis for both prosecutions were abhorrent, they seem unlikely to cause public disruption or violence. This raises the ethical concern that, without the requirement to establish this nexus, the government will have broad discretion to prosecute individuals for vivid speech that is used to express strongly held political views critical of those in power. Political speech often may be exaggerated, extreme, or offensive in making its points, but the mere fact that it takes this form should not be the basis for a prosecution. To prosecute on these grounds alone reflects the implicit assumption that the power of speech to persuade an audience to hold certain views is dangerous in itself. One alternative more consistent with liberal democratic ethical values could be to make speech that is unlikely to incite violence a lesser crime with a lesser punishment, or even a civil offence for which one may be fined. This approach, however, could still be subject to abuse by bringing actions disproportionately against those who express unpopular political opinions.

5. FREEDOM OF ASSOCIATION

The ICCPR includes the right of freedom of association (UN 1976, Article 22), as does the ECHR (Council of Europe 1952, Article 11). Both provide that a state may limit this right to prevent crime or protect national security or

public safety (Council of Europe 1952; UN 1976). Perhaps the most significant criminal measure implicating freedom of association is the prohibition of membership in a group that has been designated as a terrorist organisation. France (421-2-1), Germany (129(a)(1)), and the Netherlands (140a), for instance, prohibit membership in any organisation that has as a purpose committing a terrorist act.

More broadly, Australia criminalises not just membership in (102.3), but also association with, a terrorist organisation (102.8). Some argue that the US crime of providing material support to a terrorist organisation also has this potential. As we have described, that statute makes it unlawful to provide material support to an organisation that has been designated as an FTO, or that a person knows has engaged in terrorist activity or terrorism (18 U.S.C. §2339B (2002)).

In 2010, the US Supreme Court held that the statute did not violate the right to freedom of association when applied to a non-profit organisation that provided conflict-resolution and advocacy training to a designated FTO (*Humanitarian Law Project v. Holder*, 130 S. Ct. 2705 (2010)). The Court reasoned that, because the statute does not prohibit membership in an FTO, but the provision of support to it, the 'statute does not penalize mere association with a foreign terrorist organization' (2730). The dissent accepted the non-profit's argument that the statute should apply 'only when the defendant knows or intends that [its support] will assist the organization's unlawful terrorist actions' (2739-40). Critics argue that the dissent's position would better balance the competing interests by ensuring that only association that is intended to further terrorist activity would be a crime (Field 2014).

Moreover, it is not clear what the Court meant by 'mere association' with an FTO. Membership presumably would involve at least some activity in support of the organisation, which could then be deemed to fall within the prohibition on providing material support. In addition, the Court said that any activity that serves to enhance the perceived legitimacy of an organisation constitutes material support. This could be a basis in the future for encompassing membership within the statute, which would raise issues under the US Constitution.

As some scholars have noted, there is potential for the designation of a group as a terrorist organisation to be shaped by political considerations rather than by a rigorous analysis of the threat that it poses (Legrand 2018; Sentas 2018). While these considerations may include legitimate policy concerns, such as relationships with other states, they also may reflect efforts to impugn the legitimacy of government critics and political opposition. The fact that a state may criminalise membership in a designated organisation may be responsive to ethical concerns about vagueness and foreseeability, but that does not necessarily mean that a law is consistent with an ethical commitment to freedom of association. This underscores the need for more consensus on

what constitutes terrorism, as well as on the evidence that should be necessary in order to designate a group as a terrorist organisation.

6. FREEDOM OF MOVEMENT

Both Article 12 of the ICCPR (UN 1976) and Article 2 of the ECHR (Council of Europe 1952) guarantee a person 'lawfully within the territory of a State… the right to liberty of movement and freedom to choose his residence', and state that 'everyone shall be free to leave any country, including his own'. The right may be limited pursuant to a state's interest in 'national security or public safety', among other things (UN 1976; Council of Europe 1952).

Laws that criminalise travel to certain areas without regard to intent or activity risk violating this right. Australia, for instance, requires the prosecution simply to prove that a person entered or remained in a 'declared area' in a foreign country (Australia Criminal Code Act of 1995, 119.2(3)). The Foreign Minister designates such areas if she is 'satisfied that a listed terrorist organisation is engaging in a hostile activity in that area of the foreign country' (119.3). The defendant has the burden of proving that the travel was for a specified set of purposes, including providing humanitarian aid abroad, performing an official duty for the UN or International Committee of the Red Cross, reporting as a journalist, visiting a family member, and appearing before a court or tribunal (119.3). Only if the defendant makes a showing must the state prove that the defendant travelled with unlawful intent to engage in hostilities on behalf of a terrorist organisation (119.3).

Offences such as these have provoked considerable criticism. As one observer argues:

> it is not acceptable in a liberal democracy that a person should be [jailed] simply for traveling to an area designated by the executive branch of government as a no-go zone. It is only the fact that a person travels to that area for an illegitimate purpose that makes it worthy of criminalisation. For that reason the burden should be upon the prosecution to prove beyond a reasonable doubt that the defendant traveled for such a purpose. (McGarrity and Blackbourne 2016, pp. 142–3)

Several states prohibit travel based on an individual's intent, such as the United Kingdom (Terrorism Act 2019, Section 4), Belgium (Criminal Code Article 140sexies), and Germany (Criminal Code 89a–b). Establishing intent in such cases can be challenging (Pokalova 2020, p. 142), but the tendency of foreign fighters to use social media can help states acquire sufficient evidence of intent to participate in terrorism.

For example, the Netherlands makes it unlawful for a person to join a terrorist group abroad (Article 134a). In the first prosecution of a foreign fighter there (*Prosecutor v. Maher H.*, 1 December 2014, District Court of the Hague,

case no. 09/767116-14), Maher H., the defendant, claimed that he travelled to Syria for humanitarian purposes. The state did not have direct evidence of his activities while in Syria but presented extensive evidence to indicate that Maher H. travelled there with the intent to fight on behalf of terrorist organisations (District Court of the Hague 2014). This included his browser history, chat messages indicating indifference toward dying during jihad, a text to his mother when he was in Syria about having been on the battlefield, and photos he posted on his Facebook page with weapons and Islamic caliphate images (District Court of the Hague 2014). The state also presented videos that he had sent calling on others to participate in combat abroad (District Court of the Hague 2014). These all enabled the Netherlands to convict the defendant.

In a case such as this, the defendant's communications about his activities in Syria could support an inference that he was engaged in fighting on behalf of the terrorist group there, as could the photos on his Facebook page if they could be connected to his presence in Syria. Proving that he aided a terrorist group while in Syria based on general statements before he travelled there, however, should not be sufficient to support a prosecution. In the case of a jury trial, a court should differentiate the probative value of evidence in its instructions to a jury.

The Netherlands relies on an administrative process of issuing individualised travel restrictions to citizens, the contravention of which amounts to a criminal offence (Interim Act 2017). The Maher H. case exemplifies, however, how it could more generally criminalise and prosecute travel for the purpose of participating in terrorist activity. Moreover, it illustrates that adopting travel restrictions that require intent can still enable a state to meet its burden of proof in a criminal prosecution. This suggests that the state can respect the right to freedom of movement without unduly impairing its ability to prosecute persons who constitute genuine threats.

7. DUE PROCESS

Article 14 of the ICCPR states that 'everyone shall be entitled to a fair and public hearing by a competent, independent and impartial tribunal', and that everyone has the 'right to be presumed innocent until proved guilty' (UN 1976). Article 14(3) also codifies a right to counsel and enough time to prepare a defence, undue delay at trial, and the right to be present and defend oneself at trial (UN 1976). Article 6 of the ECHR (Council of Europe 1952) closely mirrors this provision of the ICCPR.

The Chair of the Security Council Committee on the implementation of Security Council Resolution 1373 noted some due-process concerns about prosecutions of foreign fighters in a letter to the President of the Security Council in 2016 (UN 2016, p. 115). One concern is the use of information from

intelligence agencies without disclosure to the defendant (UN 2016, p. 114). The letter emphasised the need for states to develop best practices for using such information 'without exposing sources or methods while providing for full respect for the rights of the accused' (UN 2016, p. 114). The letter also noted:

> One of the main challenges identified by the Committee's assessments is that, in some Member States, the use of 'preventive offences' is applied without full respect for several criminal law principles (e.g., the necessary precision of criminal law). The principle of legality also entails the principle of certainty (i.e., that the law is reasonably foreseeable in its application and consequences). Another criticism is that such offences may be of a 'catch-all nature' (allowing prosecutors to go 'fishing' for offences and not allowing the defence detailed knowledge of the case). (UN 2016, p. 115)

In addition, the Chair pointed out that '[q]uestions arise as to whether all family members commit an offence simply by travelling and whether they should be prosecuted even if, in some cultures, a woman must follow her husband. The question of offences committed by parents against their children by taking them to conflict zones also arises' (UN 2016, p. 116).

Two recent practices in terrorism cases also raise due-process issues. These are delegating prosecution to other states and prosecuting suspects in absentia, a practice with which France and the United Kingdom have been involved. With respect to the first, France has chosen to empower the Iraqi judiciary to handle the cases of some of its nationals accused of having fought for ISIS who were being detained abroad (Schulz 2019). France is a country that has outlawed capital punishment, but Iraq is not, and there are concerns about the integrity of the criminal justice process in Iraq (Schulz 2019). In May 2019, seven French nationals faced trial in Iraq, having admitted to playing various non-violent roles supporting ISIS in Syria (Rubin 2019). The Iraqi judge who decided the case sentenced them all to hanging (Rubin 2019). While there has not been any official determination on this issue, the practice of delegation to prosecution to another state raises the question of whether a state has human rights responsibilities to its nationals to ensure that they are not subjected abroad to treatment that would violate that state's own human rights obligations.

In addition, the Netherlands, France, and Belgium have tried terrorist defendants in absentia when those fighters remained in Iraq or Syria (Schulz 2019). States do this because they do not want to allow fighters to return to their territory, without having been charged, where they would remain free during the course of a criminal investigation (Paulussen and Pitcher 2018, pp. 22–3). This practice, however, conflicts with the right to be present and to defend oneself at trial as provided in the ICCPR and ECHR. These are funda-

mental elements of the human right to due process in criminal proceedings. One way to address state interest could be to consider whether states might be able to limit the ability of certain defendants to remain at large while under criminal investigation in a manner that would not impermissibly violate the right to liberty. Article 5 of the ECHR, for instance, permits the detention of a person 'on reasonable suspicion of having committed an offence or when it is reasonably considered necessary to prevent his committing an offence or fleeing after having done so' (ECHR Article 5(10(c))).

8. CONCLUSION: GUIDING PRINCIPLES

States have a responsibility to deter terrorism, and they have taken significant steps over the past two decades to meet this responsibility by enacting crimes relating specifically to such activity. As we have described, much of the impetus for this has come from the UN. Some have raised concern, however, that this body has de-emphasised the importance of ensuring that states are mindful of their human rights obligations when they adopt such measures. Saul (2019), for instance, argues that UN General Assembly Resolution 73/174, adopted in December 2018:

> drastically undermines the detailed human rights standards established in earlier General Assembly resolutions on counter-terrorism between 2002 and 2017. The new resolution differs from preceding ones in three key respects: (a) it omits or dilutes many earlier references to protecting specific rights when countering terrorism; (b) it is focused more on the 'detrimental effects' of terrorism on human rights than on state violations of rights by counter-terrorism measures; and (c) many provisions are not about human rights at all but, instead, are concerned with suppressing terrorism—and thus detract from a much-needed emphasis on respect for rights precisely to correct the prevailing emphasis on suppression, including from the Security Council, often at the expense of rights.

Similarly, Fionnuala Ní Aoláin (2019), UN Special Rapporteur on the Promotion and Protection of Human Rights and Fundamental Freedoms while Countering Terrorism, has observed, 'Few resources are spent to ensure that Security Council resolutions and their transposition to domestic law is compliant with human rights and rule of law'.

States should keep in mind some basic principles when considering the enactment of terrorist crimes to ensure consistency with liberal democratic values. First, when states create criminal offences linked to a domestic law definition of terrorism or terrorism acts, it is critical that the definition be clear, sufficient in scope, and precise (UN 2018, p. 38). States should avoid using broad or vague provisions to prosecute terrorists or aspiring terrorists.

The conduct proscribed by a given provision must be sufficiently clear that an individual can easily understand what behaviour is forbidden.

Second, any prosecution for expression should occur only when the state can prove that the speaker's message poses a risk to public safety, and that the speaker had such intent when disseminating the message. Overly broad prohibitions, such as apologising for or 'glorifying' terrorism, may result in criminalisation of speech that is neither intended to cause violence nor risks doing so. Such provisions also create a risk that speech will be prohibited based simply on its content, which opens the door to prosecution based on political considerations. Similarly, laws that restrict association should be drafted to require that participants have the intention to further terrorism, rather than simply because of individuals' views.

Third, criminal offences relating to travel should require that the state prove the traveller's intent to further terrorism in order to avoid undue infringement on the right to mobility. Defendants should not have to prove their innocence based on a presumption of criminal intent in order to avoid conviction, a requirement that is inconsistent with the due-process presumption of innocence in a criminal prosecution. Finally, states must abide by their obligations to ensure that defendants receive a fair trial in accordance with basic due-process standards. Any cooperation in the prosecution of a state's nationals by another state should be based on the latter state's compliance with the human rights obligations of the former state.

While liberal democracies need to combat terrorism, they also need to keep in mind that the most important advantage they can offer in the competition for influence is adhering to the values that they espouse. As Special Rapporteur Ní Aoláin (2019) has said with respect to foreign fighter laws, 'I have grave concerns that contemporary responses to foreign fighters, which have inadequately integrated human rights and humanitarian law into their regulatory scope, may in fact further inflame the "push and pull" factors that are significant to the mobilisation of young men and women to join terrorist groups'. More generally, states that disregard their human rights obligations in the quest for security risk forfeiting their claim that what distinguishes them from opponents is that they offer a way of life based on respect for human dignity.

REFERENCES

18 USC §2339A-B (2002).

Anti-Terrorism Act of 2005, Section 80.2C (Australia).

Australia Criminal Code Act of 1995, Part 5.3, Articles 119.1, 119.2(1), 119.3–4. Accessed at https://www.legislation.gov.au/Details/C2017C00235.

Chesney, Robert (2007), 'Beyond Conspiracy? Anticipatory Prosecution and the Challenge of Unaffiliated Terrorism', *Southern California Law Review* 80(3), 435.

Code Penal, Articles 140, 140sexies, 140septies (Belgium).

Code Penal, Articles 421-2-1, 421-2-5, and 421-2-6 (France).
Constitutional Council (2018), Decision no. 2018-706 QPC, 18 May 2018. Accessed at https://www.conseil-constitutionnel.fr/en/decision/2018/2018706QPC.htm.
Council of Europe (1950), European Convention on Human Rights.
Council of Europe (1952), European Convention for the Protection of Human Rights and Fundamental Freedoms, 4 November 1950. ETS 5 (entered into force 3 September 1952). Accessed at https://www.echr.coe.int/Documents/Convention_ENG.pdf.
Criminal Code of the Kingdom of the Netherlands, Article 134a (1881, amended 2012).
European Court of Human Rights (2020), Guide on Article 7 of the European Convention on Human Rights, 30 April 2020. Accessed at https://www.echr.coe.int/Documents/Guide_Art_7_ENG.pdf.
European Parliament (2017), Directive 2017/541/EU of the European Parliament and of the Council of 15 March 2017.
Field, Martha (2014), '*Holder v. Humanitarian Law Project*: Justice Breyer, Dissenting', *Harvard Law Review* 128(1), 416.
Held, Amy (2018), 'French Man Praises Police Death, Is Arrested for Apologizing for Terrorism', *NPR*, 25 March 2018. Accessed at https://www.npr.org/sections/thetwo-way/2018/03/25/596849231/french-man-praises-police-death-is-arrested-for-apologizing-for-terrorism?t=1601224359360.
Houry, Nadim (2018), 'France's Creeping Terrorism Laws Restricting Free Speech', *Just Security*, 30 May 2018. Accessed at https://www.justsecurity.org/57118/frances-creeping-terrorism-laws-restricting-free-speech.
Humanitarian Law Project v. Holder, 130 S. Ct. 2705 (2010).
Interim Act on Counterterrorism Administrative Measures (2017), Art. 8(1)–(2) (The Netherlands).
Jarvis, Lee, and Tim Legrand (2018), 'The Proscription or Listing of Terrorist Organisations: Understanding, Assessment, and International Comparisons', *Terrorism and Political Violence* 30(2), 199–215.
Legrand, Tim (2018), '"More Symbolic—More Political—Than Substantive": An Interview with James R. Clapper on the U.S. Designation of Foreign Terrorist Organisations', *Terrorism and Political Violence* 30(2), 356–72.
McGarrity, Nicola, and Jesse Blackbourne (2016), 'Anti-Terrorism Laws and Human Rights', in Leanne Weber, Elaine Fishwick, and Marinella Marmo (eds.), *Routledge International Handbook of Criminology and Human Rights*, 136, New York: Routledge.
Ní Aoláin, Fionnuala (2019), 'The Massive Perils of the Latest U.N. Resolution on Terrorism', *Just Security*, 8 July 2019. Accessed at https://www.justsecurity.org/64840/the-massive-perils-of-the-latest-u-n-resolution-on-terrorism/.
Paulussen, Christophe, and Kate Pitcher (2018), 'Prosecuting (Potential) Foreign Fighters: Legislative and Practical Challenges', ICCT Research Paper (14 January 2018), The Hague: International Centre for Counter-Terrorism.
Pokalova, Elena (2020), *Returning Islamist Foreign Fighters*. Washington, D.C.: Palgrave Macmillan.
Prosecutor v. Maher H. (2014, 1 December), District Court of the Hague Court, case no. 09/767116-14.
Resolution 1373, S.C. Res. 1373, UN Doc. S/RES/1373 (28 September 2011).
Resolution 1566, S.C. Res. 1566, UN Doc. S/RES/1566 (8 October 2004).
Resolution 2178, S.C. Res. 2178, UN Doc. S/RES/2178 (24 September 2014).

Rubin, Alissa J. (2019), 'France Hands ISIS Suspects to Iraq, Which Sentences Them to Hang', *New York Times*, 29 May 2019. Accessed at https://www.nytimes.com/2019/05/29/world/middleeast/france-iraq-isis-trials.html.

Saul, Ben (2015), 'Terrorism as a Legal Concept', in Genevieve Lennon and Clive Walker (eds.), *Routledge Handbook of Law and Terrorism*, 23 September 2015. Accessed at https://ssrn.com/abstract=2664404.

Saul, Ben (2019), 'United Nations Backslides on Human Rights in Counterterrorism', *Lawfare Blog*, 16 October 2019. Accessed at https://www.lawfareblog.com/united-nations-backslides-human-rights-counterterrorism.

Schulz, Jacob (2019), 'France Makes a Play to Try Foreign Fighters in Iraq', *Lawfare Blog*, 4 November 2019. Accessed at https://www.lawfareblog.com/france-makes-play-try-foreign-fighters-iraq.

Sections 89a and 89b, German Criminal Code of the Amendment Act on the Prosecution of the Preparation of Serious Violent Offences Endangering the State.

Sentas, Vicki (2018), 'Terrorist Organisation Proscription as Counterinsurgency in the Kurdish Conflict', *Terrorism and Political Violence* 30(2), 298–317.

Tayler, Letta (2016), 'Foreign Terrorist Fighter Laws. Human Rights Rollbacks Under UN Security Council Resolution 2178', *Human Rights Watch*, December 2016. Accessed at https://www.hrw.org/sites/default/files/news_attachments/ftf_essay_03feb2017_final_pdf.pdf.

Terrorism Act (2006), Chapter 11, Section 5 (United Kingdom).

Terrorism Act (2019), Section 4 (updating Terrorism Act 2000 581, 58B-C) (United Kingdom).

The Guardian (2015), 'Woman Jailed for "Twitter Terrorism"', *The Guardian*, 11 June 2015. Accessed at https://www.theguardian.com/uk-news/2015/jun/11/alaa-esayed-jail-twitter-terrorism-london.

The Guardian (2016), 'Security Guard Jailed for Five Years Over Tweets Glorifying ISIS', *The Guardian*, 28 April 2016. Accessed at https://www.theguardian.com/uk-news/2016/apr/28/security-guard-mohammed-moshin-ameen-jailed-for-five-years-over-tweets-glorifying-isis.

UN (1976), International Covenant on Civil and Political Rights, 19 December 1966, 999 UNTS 171, Can TS 1976 No 47 (entered into force 23 March 1976).

UN (2016), Letter dated 18 January 2016 from the Chair of the Security Council Committee established pursuant to resolution 1373 (2001), 20 January 2016. Accessed at https://www.securitycouncilreport.org/atf/cf/%7B65BFCF9B-6D27-4E9C-8CD3-CF6E4FF96FF9%7D/s_2016_49.pdf.

UN (2018), Guidance to States on Human Rights-Compliant Responses to the Threat Posed by Foreign Fighters. Accessed at https://www.un.org/sc/ctc/wp-content/uploads/2018/08/Human-Rights-Responses-to-Foreign-Fighters-web-final.pdf.

UN Security Council Counter-Terrorism Committee (2009), Model Legislative Provisions against Terrorism. Vienna: 2009. Accessed at https://www.un.org/sc/ctc/news/document/unodc-model-legislative-provisions-against-terrorism/.

2. The definition of terrorism

Seumas Miller and Jonas Feltes

1. INTRODUCTION

It is often suggested that 'one man's terrorist is another man's freedom fighter', but in fact, defining terrorism is both possible and desirable, for only then can the term 'terrorist' cease to be used purely in the context of ideological name-calling. A number of academically serious definitions of terrorism are already on offer. These definitions tend to fall into two camps. Some, such as that offered by Igor Primoratz (2013), define terrorism, in part, in terms of killing innocent persons. Others, such as that offered by Angelo Corlett (2004), define terrorism, in part, in terms of killing civilians, or at least *some category of persons that is not, by definition, or not necessarily innocent*. This issue is discussed in Section 2. A feature of most definitions of terrorism, irrespective of the camp to which they belong, is the failure to specify which of the necessary conditions that constitute the definition, including the political effects, have to be both intended and realized for the action to count as terrorism. For instance, the intentional killing of an innocent person in the service of a political purpose would normally count as an act of terrorism, but what if the action goes unreported to the public at large and, therefore, fails to have any publicity-driven political impact, although it does send its intended message to members of the security forces? This issue is discussed in Section 3. A further problem in relation to definitions of terrorism is the divide between those offered in the philosophical-ethical literature, on the one hand, and legal definitions, on the other (see Chapter 1 in this collection). Accordingly, there is a need to specify the relationship between moral and legal definitions; for example, should they simply be synonymous? This issue is discussed in Section 4.

In this chapter, a definition of terrorism is provided, but it is presented in the context of the following six assumptions (Miller 2009): (1) Terrorism is a strategy that principally consists of violent actions aimed at harming persons (directly or indirectly). Accordingly, it involves such methods as assassination (targeted killings), indiscriminate killing, torture, hostage taking, kidnapping, ethnic cleansing, and the use of chemical, biological or nuclear weapons; (2) The persons harmed are innocents or non-combatants – that is, some category

of persons of whom the deliberate killing or harming is generally regarded (albeit not by the terrorists) as an act of unjustified moral wrongdoing, for example, the deliberate bombing of a marketplace; (3) Terrorism is a means to achieve political ends (even if these are ultimately in the service of, for instance, religious ends, and even if the terrorist actions in question serve proximate ends, for example, military ends, the realization of which are a means to the political ends in question); (4) Terrorism involves terrorizing or instilling great fear in one group (typically, members of the public) in order to cause some other group (for example, their political leaders) to do what they otherwise might not have done; (5) Terrorism relies on the violent acts receiving a high degree of publicity; (6) Terrorism is a strategy that can be used by either state actors (for example, Stalin's reign of terror against his own population) or non-state actors (for example, al-Qaeda).

The first, second and third assumptions (in one form or another) feature in most definitions of terrorism. Regarding the first assumption, there is a residual issue as to whether or not violent actions directed solely at property – for example, blowing up an empty building – could count as acts of terrorism. Since this is controversial and marginal to our concerns here, we assume that such acts are not integral to the meaning of the term terrorism. Accordingly, we restrict terrorist acts to violent actions that, directly or indirectly, harm persons. Naturally, an attack on, for example, a dam that provides a city's water supply would be a terrorist attack if its intention was to indirectly harm persons by cutting off their access to water.

While the second assumption is relatively uncontroversial, the precise specification of the category of persons deliberately harmed is problematic and will receive detailed treatment in the following section.

The third assumption, that terrorism is an activity performed to realize political ends, is necessary to distinguish terrorism from other sorts of violent actions used to instil fear. Criminals, for example, sometimes use the methods of terror to achieve their criminal ends; clearly some criminals kidnap to extract a ransom, torture to instil fear and thereby extort money, and so on. On the other hand, it needs to be noted that sometimes terrorism has multiple ends; for example, the Islamic State's (ISIS) terrorist methods in theatres of war in Iraq and Syria serve military, political and religious ends.

The fourth assumption is essentially a conceptual claim. For an activity to count as terrorism, someone has to be trying to terrorize someone else, and for terrorism to be a strategic activity – as opposed to, for example, merely an expressive activity – it has to be in the service of some further end, that is, changing the attitudes and/or behaviour of some group.

The argument for the fifth assumption is as follows. If fear is to be instilled in some group – for example, members of a community – as a consequence of the harm done to some other group or subgroup – for example, victims of

bombings – then the first group needs to know that the second group has in fact been harmed. Accordingly, the terrorist strategy relies on a high degree of publicity. Indeed, it might well be that, other things being equal, the higher the level of publicity, the more successful the terrorist strategy is. This certainly was the case with al-Qaeda's terrorist attack on the Twin Towers on 11 September 2001.

The sixth assumption is controversial. Some definitions, particularly those offered by nation states – for example, the United States – restrict terrorism to non-state actors. However, terrorism is a strategy that is available to both state and non-state actors; indeed, historically, it is a strategy that has been used on a larger scale by state actors than by non-state actors (Primoratz 2013, Ch. 2). Accordingly, the decision to exclude state actors from the definition is either gratuitous or, more likely, based on political motives, for example, a desire on the part of nation states not to implicate themselves.

In light of these six assumptions, we offer the following preliminary definition of terrorism.[1] By definition, terrorism is a strategy that (Miller 2005, 2009, Ch. 2):

1. Consists of state or non-state actors deliberately perpetrating acts of violence aimed at (directly or indirectly) seriously harming persons, the deliberate harming of whom is (other things being equal) generally regarded as morally wrong, for example, children;
2. Is a means of terrorizing the members of some social, economic, political, ethnic or other group to achieve a political end; and
3. Relies on the violence receiving a degree of publicity, at least to the extent necessary to engender widespread fear in the target group.

2. TERRORIST TARGETS: INNOCENTS AND CIVILIANS

The first condition in our preliminary definition of terrorism is problematic in that it does not sufficiently specify the category of persons against whom violence is deliberately used. However, as mentioned above, definitions typically specify the category of persons in question as either being innocents or as being civilians, that is, non-combatants. Let us consider each of these two types of definition beginning with the one related to innocents.

This type of definition is open to counterexamples (Miller 2009). Imagine a non-democratic, indeed highly authoritarian, government pursuing policies that are widening the gap between the rich and the poor. Assume that well-intentioned democrats with a social conscience attempt to mobilize opposition to the government – opposition in the form of non-violent protests, strikes, boycotts, dissemination of anti-government material, passive

non-compliance, and so on. These opposition elements are seeking to overthrow the government, indeed the system of government, albeit by non-violent means. The African National Congress (ANC) in its initial *non-violent* phase prior to the 1960s is a case in point. Accordingly, they are not innocents in the required sense (indeed, from the perspective of the authoritarian government, these opposition forces are engaged in attempting to overthrow the *legitimate* government of the country). Moreover, they may well succeed if harsh countermeasures are not introduced. Accordingly, the government embarks on a campaign of killings ('disappearances') and torture of opposition elements to instil fear in the opposition forces as a whole, and thus put an end to the 'insurrection'. Surely this is state terrorism of the kind practised by the Argentinian generals in the 1980s and (to a lesser extent) by the apartheid government in South Africa against the ANC. Nevertheless, it remains the case that the opposition forces are responsible for attempting to overthrow the government, and the government believes itself – and is believed by many, let us assume – to be legitimate. Based on the definition of terrorism in terms of innocents, the killings and torture perpetrated by the government are not terrorism since the opposition forces are not innocent in the required sense.

What of the definition of terrorism in terms of civilians, that is, non-combatants (Miller 2007, 2009)? Consider corrupt senior government officials and civil servants who fail to organize the distribution of aid in the form of medicine and food to their starving, disease-afflicted fellow citizens, but rather sell it to line their own pockets. Suppose the foreseen consequence of this corruption and dereliction of their humanitarian duty is that tens of thousands of the needy die. These officials are not combatants in the required sense; they are not, themselves, soldiers engaged in an armed attack, nor are they the leaders of such combatants or assisting such combatants qua combatants. Accordingly, targeting these public officials would be, according to the definition before us, terrorism. But these officials are guilty in the sense that they are morally responsible for ongoing, widespread and serious rights violations. Moreover, using lethal force against some such officials to instil fear in their fellow guilty officials, and thereby bring about a cessation to these ongoing, widespread and serious rights violations, may well be, under certain circumstances, morally justifiable. It seems that such actions should be regarded as protection measures against rights violations rather than terrorism.

In light of these counterarguments to the definitions of terrorism in terms of violence directed at innocents and non-combatants, how do things now stand?

Let us begin by making the point that violence directed at *military* combatants (including the leaders of military combatants) in theatres of war is not terrorism even if it otherwise meets our definition of terrorism. This thought is one of those motivating the definition of terrorism in terms of targeting non-combatants. Secondly, violence directed at state or non-state actors who

are perpetrating serious, ongoing and widespread human rights violations is not necessarily terrorism – for example, the ANC's early-1960s switch to the use of violence against apartheid state actors who were engaged in ethnic cleansing (forcible removals), torture of activists and so on. While there is obviously a grey area here in relation to human rights violations, we can distinguish between, on the one hand, human rights violations at the extreme end of the scale and perpetrated on a large scale, (for example, genocide, ethnic cleansing, mass starvation) and, on the other hand, curtailments of civil and political rights and perpetrating social and economic injustices that stop short of human rights violations at the extreme end (for example, inequalities of wealth and opportunity). Thirdly, state actors who use violence against *violent* revolutionary non-state actors are not necessarily terrorists, even though the violence of these state actors might meet the other conditions of our definition. Indeed, some violent revolutionary non-state actors are de facto military combatants, for example, ISIS.

Accordingly, we suggest that terrorists direct violence at persons who are *not* military combatants, human rights violators (perpetrating large scale, ongoing, serious human rights violations) or violent revolutionaries. Therefore, our definition of terrorism becomes (Miller 2005, 2009, Ch. 2):

Terrorism is a strategy that:

1. Consists of state or non-state actors deliberately perpetrating acts of violence aimed at (directly or indirectly) seriously harming persons who are *not* military combatants, human rights violators or violent revolutionaries;
2. Is a means of terrorizing the members of some social, economic, ethnic, political or other group to achieve a political purpose; and
3. Relies on the violence receiving a degree of publicity, at least to the extent necessary to engender widespread fear in the target group.

3. INTENTIONS OF TERRORISTS[2]

An issue or, rather, set of issues that now arises concerns the intentions of the terrorists; specifically, do all of their intentions need to be realized for their actions to count as instances of terrorism? Here, there are three main intentions of interest: the intention to use violence against persons; the intention to create widespread fear (and to do so relying, in part, on the violent act and harm done being made public); and the intention to achieve some political purpose. Accordingly, three corresponding questions arise. Is it necessary to actually perform an act of violence against a person, for example, against an innocent non-combatant? Is it necessary to actually instil fear in the target audience? Is it necessary to achieve the political outcome aimed at? Let us discuss these three questions in order.

Presumably, a group of would-be terrorists who fail to perform their intended act of violence because, for instance, the bomb they planted fails to detonate have not, thereby, performed the terrorist act in question; rather, they have merely *attempted* to do so (and failed in that attempt). The would-be terrorists have not performed a violent act, and the violent act *is* the terrorist act. Naturally, even attempted terrorist actions could be criminalized (as, for instance, is attempted murder), but that is a different matter. Attempted murder is not murder, and attempted terrorism is not terrorism.

What of the intention to instil fear in the target audience (relying, in part, on the violent act and harm done being made public)?[3] Does this intention need to be realized for an act of violence to count as a terrorist act? Consider the following example.

On 17 October 2015, the German right-wing extremist Frank S. attacked the candidate for mayor of Cologne Henriette Reker at a rally in Cologne Braunsfeld with a bowie knife. After stabbing the politician in the neck, S. assaulted and wounded four bystanders (Rath 2015; *The Irish Times* 2015). The assault was stopped by German federal police officers. After his arrest, the attacker repeatedly named the refugee-friendly policies of Reker, German chancellor Angela Merkel, and other German politicians as a motive for the attack. During the trial against S., the German Federal Prosecutor General characterized the attack as 'intended to create a climate of fear among all persons engaged with refugee affairs'.[4] S. was sentenced to 14 years in prison for attempted murder and grievous bodily harm in four cases (Deutsche Welle 2016).

However, although clearly intended by S., the attack did not create widespread fear in society. Reker was elected mayor of Cologne only one day later while still in a coma, and her political opponent, Jochen Ott, stopped his campaign on 17 October out of solidarity (Rath 2015). Furthermore, because Frank S. was arrested during the attack and was clearly identified as a lone operator, the citizens of Cologne did not expect further attacks. Not fear, but anger and outrage, dominated the public discourse after the attempted assassination of Henriette Reker. Thus, Frank. S. committed an act of terrorism. Moreover, in performing his act of terrorism, he intended to cause widespread fear but failed to do so.

The example also serves to demonstrate that the political end that a putative terrorist act is intended to serve does not have to be realized for the act to constitute an act of terrorism. After all, Frank S.'s act clearly failed to achieve its political purpose and, indeed, might have strengthened the political forces he had hoped to diminish.

In concluding this section, we need to briefly mention the view that intentions are not necessary for acts to count as acts of terrorism. This view is surely false if it implies that an act of terrorism could be an act that was not

intended to terrorize and had no political end. However, some theorists (Kamm 2011, pp. 73–118; Rodin 2004, pp. 752–71) evidently hold that a violent act that caused *unintended harm* to a person or persons and was performed by members of a group who intended, in performing this act, to terrorize in the service of a political agenda might count as an act of terrorism. Consider the following example.

On the night of 23 August 1970, an explosive device detonated behind Sterling Hall at the University of Wisconsin in Madison killed the physicist Robert Fassnacht (Cronin and Jenkins 1999, p. 517). The perpetrators of this attack were later identified as Dwight Armstrong, his brother Karleton Armstrong, David Sylvan Fine, and Leo Burt. The Armstrong brothers planned and executed the attack together with their co-conspirators as members of the radical left-wing group the 'New Year's Gang' (Cronin and Jenkins 1999, p. 517). According to the group, no civilians should have been hurt in the attack that was aimed at the Army Mathematics Research Center in Sterling Hall (New Year's Gang 1970, p. 1). However, although the group executed a warning call, the detonation occurred prematurely and thereby killed Fassnacht, who happened to be in the building at that time (Bates 1993, p. 307; Fellner 1986).

This example is an interesting borderline case between sabotage and terrorism. Intuitively, many people would call – and have called – the New Year's Gang a terrorist group despite the fact that the group *did not intend* to harm persons, but rather merely to damage buildings, to further their political aims. Evidently, the members of the group intended to promote their political aims via a well-publicized, fear-inducing act of violence. So the example meets all our conditions for a terrorist act other than the intention to harm persons. But was it an act of terrorism or merely a politically motivated act of sabotage that went wrong?

Shortly after the attack, the news media reported that the New Year's Gang issued a warning call prior to the attack to avoid casualties. Moreover, when they claimed responsibility for the attack, the group also expressed regret over Fassnacht's death (New Year's Gang 1970, p. 1). While this latter piece of information did not extinguish the fear, anger and moral outrage felt by members of the community, it did mitigate, in particular, the fear that they could well be the targets of further attacks and that, as a consequence, further lives may be lost.[5] Moreover, the fact that the killing of Fassnacht was *unintentional* was also the basis for the subsequent indictments against the group characterizing their crime as an act of sabotage with manslaughter (third-degree murder) – but not as terrorism. Accordingly, we conclude that intention to seriously harm persons is, after all, a necessary condition for an act to qualify as terrorism.

In light of this discussion, our definition of terrorism becomes:

Terrorism is a strategy that:

1. Consists of state or non-state actors deliberately performing acts of violence aimed at (directly or indirectly) seriously harming persons who are *not* military combatants, human rights violators or violent revolutionaries;
2. Is an intended means of terrorizing the members of some social, economic, ethnic, political or other group to achieve a political purpose; and
3. Relies on the violence receiving a degree of publicity, at least to the extent necessary to engender widespread fear in the target group.

4. TERRORIST ACTIONS: MORALITY AND LAW

At this stage of proceedings, our definition demarcates many, if not most, terrorist actions from both non-violent actions, and from violent actions that are not terrorist actions. Unfortunately, the definition is still incomplete by virtue of leaving a degree of indeterminacy, including in relation to legitimate types of violent attacks and also in relation to legitimate targets of violent attacks, for example, specification of the category of human rights violators. However, this is to be expected if we grant, as it seems we must, that the concept of terrorism is somewhat vague. Moreover, it has the consequence that there is some room for us to be stipulative in relation to types and targets of violent acts, in particular.

Granted that there is this room for stipulation, we need to determine what purposes would be served by this or that stipulative definition of terrorism (or definitional element thereof). We suggest that an important purpose in defining terrorism is to render it a serious crime – a serious crime both in terms of domestic and international law. Here, we are assuming that the notion of crime in play is (at least) that of a serious form of moral wrongdoing, objectively considered (obviously, crime is also a form of unlawful action). So murder is a serious crime, but shoplifting typically is not, and neither are homosexual acts between consenting adults. Shoplifting is not a sufficiently serious form of moral wrongdoing to count as a serious crime, and homosexuality fails the test of objectivity (albeit some people believe it is a serious form of moral wrongdoing).

However, we need to keep in mind that there is a distinction between the concept of a serious crime and the concept of a morally justifiable act. Accordingly, there is the conceptual possibility of some action being both a serious crime and being morally justifiable. Thus, torture is a serious crime; however, arguably, torture might be morally justifiable in some extreme circumstances. The point is that defining terrorism in such a way as to render it a serious crime (or at least an act that ought to be a serious crime) does not settle the question as to whether or not it is morally justifiable (at least in all

circumstances). Naturally, since criminal law tracks morality, the fact that some kind of act – for example, murder or torture – is a serious crime implies that *in general* – indeed, in all but the most extreme circumstances – it is morally unjustified.

The suggestion, then, is that we should further demarcate terrorist actions by insisting that they are violent acts that are or, more precisely, should be criminalized. Accordingly, as a preliminary, we should trawl through the statute books, human rights charters, and so on of *relevant* jurisdictions and identify the justifiably accepted – and de facto more or less universally accepted – set of serious violent crimes against the person, such as murder, torture, grievous bodily harm, rape and kidnapping (jurisdictions that are not relevant would include totalitarian states and other nation states that are beyond the pale).

This initial *long* list of existing serious violent crimes that are justifiably serious crimes is then cross-tabulated with our set of defining features and additional criteria of terrorist actions to generate a new (shorter) list of violent actions. This shorter list constitutes our initial set of terrorist actions; however, it should be added to if and when other violent crimes are justifiably legislated against as violent crimes, and meet the other criteria for being terrorist actions. Accordingly, we recommend that our above definition of terrorism be augmented by a fourth condition, namely, that the violent actions in question be ones that ought to be criminalized.

In light of the discussion in this section, our definition of terrorism becomes (Miller 2005, 2009):

Terrorism is a strategy that:

1. Consists of state or non-state actors deliberately performing acts of violence aimed at (directly or indirectly) seriously harming persons who are *not* military combatants, human rights violators or violent revolutionaries;
2. Consists of violent actions that ought to be criminalized;
3. Is an intended means of terrorizing the members of some social, economic, ethnic, political or other group to achieve a political purpose; and
4. Relies on the violence receiving a degree of publicity, at least to the extent necessary to engender widespread fear in the target group.

A final point in relation to the above definition arises from a consideration of the purposes of defining terrorism and, in particular, the purposes of legal definitions of terrorism. Accordingly, this point pertains to condition (2) above. One important purpose of criminalizing acts of terrorism is to combat terrorism. Hence, there are many terrorism laws that criminalize assisting terrorists, for example, by way of financing terrorist groups or training terrorists. This raises the question of the limits that ought to be placed on such laws. In liberal democracies, these limits are, in part, to be determined by recourse to individ-

ual rights, especially those pertaining to various freedoms, such as freedom of expression, of movement and so on (see Chapter 1 for a discussion of these issues). However, these complex matters cannot be pursued here.

NOTES

1. Earlier versions of these definitions of terrorism appeared in Miller (2009, Ch. 2).
2. This section is derived from Feltes's (2020) PhD submitted to the Delft University of Technology and titled 'CBRN Threats, Counter-Terrorism and Collective Responsibility'.
3. We assume public knowledge is intended since fear of something requires belief in at least its potential existence.
4. Translated from the original German: 'S. habe ein "Klima der Angst" bei allen in der Flüchtlingsunterbringung engagierten Personen erzeugen wollen'. See Rath (2015).
5. Yet, admittedly people might have been scared to a certain degree to fall victim to another failed act of sabotage by being in the wrong place at the wrong time.

REFERENCES

Bates, Tom (1993), *Rads: The 1970 Bombing of the Army Math Research Center at the University of Wisconsin and Its Aftermath*, New York: HarperCollins.

Corlett, Angelo (2004), *Terrorism: A Philosophical Analysis*, Dordrecht: Kluwer.

Cronin, David & Jenkins, John (1999), *The University of Wisconsin. A History, 1945–1971 Renewal to Revolution*, Volume IV, Madison: University of Wisconsin Press.

Deutsche Welle (2016), 'Man who stabbed mayor of Cologne sentenced to 14 years in jail', *DW Online*, accessed 10 August 2020 at: http://www.dw.com/en/man-who-stabbed-mayor-of-cologne-sentenced-to-14-years-in-jail/a-19371698.

Fellner, Michael (1986, May 18), 'The Untold Story. After 15 Years, Dwight and Karl Armstrong Reveal the Drama Behind the Anti-Vietnam War Bombings in Madison, Part II', *The Milwaukee Journal*.

Feltes, Jonas (2020), *CBRN Threats, Counter-Terrorism and Collective Responsibility*, PhD dissertation submitted to Delft University of Technology, the Netherlands.

Kamm, Frances M. (2011), *Ethics for Enemies: Terror, Torture, and War*, Oxford: Oxford University Press.

Miller, Seumas (2005), 'Terrorism and Collective Responsibility', in: G. Meggle (ed.), *Ethics of Terrorism and Counter-Terrorism*, Frankfurt: Ontos Verlag.

Miller, Seumas (2007), 'Civilian Immunity, Forcing the Choice and Collective Responsibility', in: I. Primoratz (ed.), *Civilian Immunity*, Oxford: Oxford University Press.

Miller, Seumas (2009), *Terrorism and Counter-Terrorism: Ethics and Liberal Democracy*, Oxford: Blackwell.

New Year's Gang (1970), 'Statement Concerning UW bombing', *Kaleidoscope* Vol. 2/17 (August 30, 1970).

Primoratz, Igor (2013), *Terrorism: A Philosophical Investigation*, Cambridge: Polity Press.

Rath, Christian (2015), 'Ermittlungen nach Reker-Attentat: "Gefahr für das Ansehen der Bundesrepublik im Ausland"', *Koelner Stadtanzeiger*, accessed 10 August 2020 at: http://www.ksta.de/koeln/gefahr-fuer-das-ansehen-der-bunesrepublik-im-ausland-sote-23078416.

Rodin, David (2004) 'Terrorism without intention', *Ethics*, 114(4), 752–71.

The Irish Times (2015), 'Cologne Mayor candidate stabbed and severely injured–police', accessed 10 August 2020 at: http://www.irishtimes.com/news/world/europe/cologne-mayor-candidate-stabbed-and-severely-injured-police-1.2396010.

Whittaker, David J. (ed.) (2003), *The Terrorism Reader*, 2nd edition, London: Routledge.

3. Collective responsibility and counter-terrorism

Seumas Miller and Jonas Feltes

1. INTRODUCTION

A terrorist attack of any significance, such as the 9/11 attack on the Twin Towers in New York and other sites in the US, is rarely committed by a lone actor but rather involves the actions of multiple actors who are working towards a common goal. Thus, such a terrorist attack can be characterized as a joint action performed by the members of a group of actors – in this case, terrorists. Roughly speaking, a joint action is an action comprised of a set of individual actions, each of which is directed to the same end (a collective end, in our parlance). Thus, two men lifting a crate onto a truck is a joint action; each lifts his side of the box and, in doing so, each has as an end to relocate the crate from the ground onto the truck. Moreover, each does his part believing that the other will do his part; there is interdependence of action. However, some joint actions, such as a large number of workers building, for example, the Great Wall of China or soldiers fighting a war, are far more complex and take place over a far longer period of time.[1]

What of those combating terrorist attacks? The countermeasures against terrorist attacks also should be regarded as the joint actions of multiple members of various groups of actors having as their collective end to prevent or respond to these attacks.

However, whenever agents cooperate to realize a collective end, questions of responsibility arise – not only with regard to causal responsibility but especially concerning *moral* responsibility. Here, it is necessary to take a step back and investigate the following questions: Who is (morally) responsible for the consequences that result from an action performed by members of a group of actors? Can a group itself be held responsible for a specific action? Or is it only the members of a group that are the bearers of responsibility for a joint action to which they contribute? On the view defended here, it is the latter: only individual persons, and not collective entities per se, can properly be held morally responsible for actions (joint or otherwise). The second section

of this chapter provides an analysis of the notion of the moral responsibility of members of groups who perform joint actions. According to this analysis, the so-called collective responsibility of a group is to be understood in terms of the joint responsibility of the individual persons who comprise that group.[2] In the third section, we apply this notion of joint responsibility to terrorist groups that perform a terrorist attack and to the members of security agencies and others who would seek to prevent or respond to such an attack.

2. COLLECTIVE MORAL RESPONSIBILITY

There are at least four senses of responsibility that apply to single individuals and groups of individuals alike. A single person who intentionally performs an action (and whose intention is under their own control) is responsible for that action.[3] However, since many actions are not morally significant, this notion of responsibility is not yet moral responsibility. Let us refer to this first sense of responsibility as *natural* responsibility and, in the case of a single person's action, *individual* natural responsibility. Likewise, persons who perform a joint action are responsible for that action in the sense of natural responsibility. However, since a joint action is performed by multiple persons, this is *collective* natural responsibility. Accordingly, to say that they are collectively responsible for the action is just to say that they performed the joint action. That is, they each had a collective end, each intentionally performed their contributory action, and each did so because each believed the other would perform their contributory action, and that therefore the collective end would be realized.

Our second sense of responsibility is *institutional* responsibility. If a role occupant has an institutionally determined obligation to perform an action – for example, a police officer might have an obligation to make an arrest – then the person is (individually) *institutionally* responsible for making the arrest. Likewise, if the occupants of an institutional role (or roles) have an institutionally determined obligation to perform some joint action, then those individuals are *collectively* institutionally responsible for its performance, in our second sense of collective responsibility. Here, there is a *joint* institutional obligation to realize the collective end of the joint action in question. In addition, there is a set of derived *individual* obligations; each of the participating individuals has an individual obligation to perform their contributory action (the derivation of these individual obligations relies on the fact that, if each performs their contributory action, then it is probable that the collective end will be realized).

The third sense of responsibility is *moral* responsibility. If a person intentionally performs a morally significant action, then they are (individually) *morally* responsible for the action.[4] Likewise, if a group of persons performs a morally significant joint action, then they are *collectively* morally respon-

sible for it. Moreover, as suggested above, collective moral responsibility for outcomes that are intended, or otherwise aimed at, is a species of *joint* responsibility. Accordingly, each agent is individually morally responsible, but conditionally on the others being individually morally responsible; this interdependence in respect of moral responsibility exists because the action of each is performed in the service of a collective end.

Thus, we can make the following claim about collective moral responsibility: if multiple persons are collectively (that is, jointly) – naturally or institutionally – responsible for the realization of an end (an outcome), and if the end, and therefore outcome, is morally significant, then – other things being equal – the persons are collectively (that is, jointly) morally responsible for that outcome, and can reasonably attract moral praise or blame, and (possibly) punishment or reward, for bringing about the outcome.

Here, we need to be more precise about what persons who perform morally significant joint actions are collectively morally responsible for. Other things being equal, each person who intentionally performs a morally significant *individual* action has *individual* moral responsibility for the action. So, in the case of a morally significant joint action, each person is *individually* morally responsible for performing *their contributory* action, and the *other* persons are *not* morally responsible for their individual contributory action. In addition, however, the contributing persons are *collectively* morally responsible for the outcome or *collective end* of their various contributory actions. To say that they are collectively morally responsible for bringing about this (collective) end is just to say that they are *jointly* morally responsible for it. Thus, each person is individually morally responsible for realizing this (collective) end, but conditionally on the others being individually morally responsible for realizing it as well.

2.1 Layered Structures of Joint Action

In our discussion above, we distinguished between natural, moral and institutional responsibility and, more specifically, between collective natural, collective institutional and collective moral responsibility. Let us now focus attention on collective institutional responsibility in particular. For our purposes here, an institution can be understood as an organization or system of organizations constituted at least in part by a structure of roles and by some collective end(s) served by that structure of roles (Miller 2010).[5] For instance, a military organization fighting a battle might consist of officers, infantry soldiers, tank crews, pilots and so on, and have as a collective end to win the battle. Notice that the joint actions performed by the occupants of such organizations often consist of *layered structures of joint actions* (Miller 1992, pp. 275–97; 2001, Ch. 5). For instance, the members of the organization's infantry platoon might have as

their collective end to take and hold the ground occupied by the enemy (joint action j_1), the members of the tank crews might have as their collective end to destroy the enemy gun emplacements (joint action j_2), and the pilots comprising the squadron might have as their collective end providing air cover for the infantry and tanks (joint action j_3).

Let us refer to the large-scale, complex joint action that consists in winning the battle as *J*. *J* consists of the actions of all the above – that is, infantry, tank crews and pilots. Moreover, *J* consists in the subsidiary joint actions, j_1, j_2 and j_3; the collective ends of each of these subsidiary joint actions – for example, to take and hold ground – ultimately serves the collective end of *J*, that is, to win the battle. Other things being equal, we can now say that all or most of the members of the above military units have, at least in principle, collective responsibility – that is, joint natural responsibility – for winning the battle (supposing they do win it) by way of their participation in a layered structure of joint actions. Of course, things might not be equal if, for instance and as mentioned above, many of these persons did not perform their actions having as at least one of their ends to win the battle, but rather, for instance, to simply avoid being shot for desertion.

However, institutional role occupants have more than simply natural responsibility (individual or joint) for their actions and omissions. Institutional role occupants are governed by sanction-backed regulations and laws that both constrain and enable the actions that they (institutionally, for example, legally) ought, and ought not, to perform qua institutional role occupants (for example, in the case of a military organization, the laws of war). If the occupants of institutional roles have institutional responsibilities with respect to their performance of joint actions (or joint omissions), then these responsibilities are collective institutional responsibilities. Note that in some cases these collective institutional responsibilities will be prospective, such as in cases where there is a *joint* institutional duty to realize the collective end of some joint action. Here, the individual duty of each to perform their contributory action is interdependent with the individual duty of each of the others to perform theirs. On the other hand, as was mentioned above, collective institutional responsibility can also be retrospective, such as in cases where the institutional actors have failed to do their joint duty. Note also that, while institutional responsibilities are often congruent with moral responsibilities, this is not necessarily the case. In apartheid South Africa, police were legally, that is, institutionally, required to enforce morally repugnant laws and policies, such as the Group Areas Act and the forcible removal of blacks to desolate so-called homeland areas. Such lawful actions resulted in the armed struggle of the African National Congress and, in particular, of its armed wing, Umkhonto we Sizwe (Spear of the Nation).

2.2 Chains of Responsibility

Let us now turn to the application of our theory of collective responsibility as joint responsibility to morally significant *diachronic* institutional action. Consider a team of detectives investigating a terrorist bombing. Let us assume that the team is engaging in a joint institutional action, namely, that of determining the identities of the terrorists. Members of the team gather physical evidence and interview witnesses and, in particular, any suspects. Moreover, they do so having as a collective end to determine the *factual* guilt or innocence of these suspects. At some point the detectives complete this process and provide a brief of evidence to the prosecutors according to which, and based on all the evidence, certain identified individuals perpetrated the terrorist bombing. So far so good, but the criminal justice processes do not terminate in the work of the detectives for there is now the matter of the trial; that is, the determining by the members of a jury of the legal guilt or innocence of the suspects. Let us assume that the members of the jury perform the joint (epistemic) action (Miller 2018a, pp. 300–318) of deliberating on the *legal* guilt or innocence of the suspects, and jointly reach the verdict of guilty. The question that now arises concerns the institutional relationship between the joint institutional action of the detectives and the joint institutional action of the members of the jury. It is here that the notion of a chain of institutional responsibility is illuminating (Miller 2014, pp. 21–39).

Let us assume in what follows that the collective end of the criminal justice process comprised of both the investigating detectives *and* the members of the jury (as well as others, but here we simplify) is that the factually guilty be found legally guilty (and the factually innocent not be found legally guilty). Note that from the perspective of this larger institutional process, the collective end of the detectives (that of determining the factual guilt or innocence of a suspect) is merely *proximate* whereas that of the members of the jury is *ultimate* (it is, of course, only penultimate from the perspective of the criminal justice system more broadly conceived, given the need for sentencing and incarceration).

Moreover, in all this there is an institutional division of labour and segregation of roles that involves each type of institutional actor – for example, investigator, prosecutor, judge, jury and others – making a contribution to the further (collective) end of identifying and appropriately punishing the guilty and exonerating the innocent. However, unlike many institutional arrangements, the criminal justice process is predicated on strict adherence on the part of institutional actors to the segregation of roles on pain of compromising this further end. We emphasize that this segregation of roles is consistent with all of these actors, each with their own different and segregated role, having a common further aim; agents can have a common aim and yet it is a requirement that each is to make a different and distinct contribution to that aim, and

not perform the tasks assigned to the others, and do all this in the service of that common aim.

3. TERRORISM AND COUNTER-TERRORISM

3.1 Collective Responsibility of Terrorists

Whatever the political aims of terrorist groups, and these are multiple and – presumably at least in some cases (for example, those directed at colonial powers) – morally worthy, their methods typically (if not by definition) comprise the murder of innocent persons, including children (in the case of extremist jihadist groups, such as al-Qaeda and the Islamic State), and therefore are morally objectionable. The possibility of morally worthy ends being pursued by morally objectionable means can give rise to moral dilemmas; do the ends justify the means? Accordingly, there is at least the notional possibility that some terrorist attacks are morally justified. That said, in the case of extremist jihadist groups, such as Islamic State, there is no moral dilemma. First, the end of establishing an authoritarian, indeed fascist, state (the so-called caliphate) in which human rights (for example, those of women and unbelievers) are violated is morally unacceptable. Second, the means to that end include large-scale atrocities, such as genocide and enslavement, for example, against the Yazidis (Spencer 2014) is morally unacceptable.

The preparedness of members of the Islamic State, al-Qaeda and other extremist jihadist groups to commit suicide, and thereby supposedly achieve martyrdom, is an enormous advantage for a terrorist organization. Moreover, this role is greatly facilitated not only by real and perceived injustices, and existing national, ethnic and religious conflict, but also by global financial interdependence and modern technology, such as the global communication system and the new chemical and biological weapons of mass destruction that these groups have been seeking to develop. Perhaps al-Qaeda's success is not dependent on widespread political and popular support for its goals, although it is certainly reliant on disaffection, including with US policies. Rather, its success might largely be a function of the psychological preparedness and logistical capacity to perpetrate acts of terror, coupled with the technological capacity to communicate those acts worldwide, and thereby wreak havoc in a globally economically interdependent world. Its methods have proved extraordinarily effective in relation to the goal of destabilization. The terrorist group from the medieval past has identified the Achilles heel of the modern civilized world.

At any rate, from the perspective of this article, the members of the Islamic State and al-Qaeda bear collective *natural* responsibility for these various attacks and their intended (and perhaps foreseeable) outcomes, and, since these

attacks are clearly morally significant – indeed, morally blameworthy – the members of these groups are collectively morally responsible – indeed, collectively morally culpable – for these attacks. They are morally culpable because, as already mentioned, their methods clearly involve the intentional killing of the innocent, and are not constrained by principles of the proportional use of force or minimally necessary force. Indeed, the collective end of people like Osama bin Laden, Abu Bakr al-Baghdadi (the former leader of the Islamic State) and their followers and successors has been to maximize the loss of human life (albeit apparently in the service of their ultimate collective end of establishing a caliphate and so on). It remains an open question whether this is so for *all* forms of terrorism.

It is obvious that terrorist attacks are typically joint actions and, therefore, in light of our discussion in Section 2, the perpetrators of these attacks are collectively, that is, jointly, morally responsible for these attacks and the murder of the victims of the attack. For example, the terrorists who hijacked American Airlines Flight 11 and crashed the plane into the North Tower of the World Trade Center in New York performed a joint action. At least one terrorist operated the controls of the plane, while another navigated, and the remaining terrorists, by violence and the threat of violence, prevented the cabin crew and passengers from intervening. Each performed a contributory action, or actions, in the service of the collective end of crashing the plane into the building and killing passengers, office workers and themselves. Accordingly, the terrorists are collectively, that is, jointly, morally responsible for the murder of the passengers and of the occupants of the World Trade Center.

Further, since these members of these terrorist groups perform tasks as members of organizations (even if, to some extent, loosely organized organizations), the notion of a layered structure of joint action becomes relevant. Thus, the Islamic State's successful attack on the city of Mosul in Iraq was a manifestation of a layered structure of joint action (see Section 2 above). This is because it was a complex cooperative enterprise and, therefore, those who participated in it can, *at least in principle*, be ascribed collective, that is, joint, *natural* responsibility for the outcomes aimed at, and in fact realized, in undertaking that enterprise. Moreover, since the enterprise was morally significant, they can also be ascribed collective, that is, joint, *moral* responsibility for these outcomes. Note that such structures involve: (1) a possibly indirect and minor causal contribution from each of the individuals jointly being ascribed responsibility; (2) each individual having an intention to perform their contributory (causally efficacious) action; (3) each individual having as an ultimate end or goal the outcome causally produced by their jointly performed actions; (4) some individuals – for example, those holding leadership roles – having a greater degree of moral responsibility than others; and (5) some having

diminished moral responsibility by virtue of, for instance, being coerced into participating.

Naturally, here, as elsewhere, important questions arise in relation to those who assist terrorist organizations without being members of them – for example, providers of financial assistance – or who act in their name without being members – for example, some 'lone-wolf' terrorists.[6] These latter actors may or may not fall within the ambit of a layered structure of joint actions and, therefore, are outside the reach of the collective responsibility therefrom derived. However, even if they do not, they are likely to be able to properly be ascribed moral responsibility, indeed moral culpability, for their actions and, therefore, justifiably be investigated, tried and punished as criminals.

3.2 The Web of Prevention

As mentioned above, the investigation of terrorist attacks and the like typically involves joint action on the part of institutional actors, such as police, and, therefore, collective, that is, joint, institutional and moral responsibility, including in the context of a chain of responsibility. However, counter-terrorism (CT) writ large, so to speak, involves cooperation between multiple security and other state agencies, financial institutions and other businesses, and members of the public. Indeed, it involves what has in other contexts been termed a 'web of prevention'. As such, it involves multiple, coordinated, layered structures of joint action, that is, *iterated* layered structures of joint action. Moreover, since the web of prevention also has a diachronic dimension that consists of the operation of institutional processes in which multiple institutional actors function in accordance with a division of labour, it involves complex, intersecting chains of responsibility.

The concept of a so-called 'web of prevention' is based on the notion of collective action and collective responsibility (here understood as joint responsibility) and was initially introduced in the domain of biosecurity. The concept was originally mentioned in an initiative of the International Committee of the Red Cross on biotechnology and security in 2002 (Rappert & McLeish 2012, p. 4; Selgelid & Rappert 2013, p. 277). Yet, similar concepts, such as the web of deterrence, date back to debates of non-proliferation and biosecurity during the Cold War (Rappert & McLeish 2012, pp. 3–4).

In the context of biosecurity, the concept of the web of prevention describes an 'integrative and comprehensive approach' (Whitby et al. 2015, Ch. 7) to prevent the malicious use of biotechnology as weapons. The web involves a variety of stakeholders, such as national security institutions, international organizations as well as research institutions. These groups of stakeholders are jointly responsible for implementing a set of integrated measures, such as export controls, disease detection and prevention, effective threat intelligence,

international and national prohibitions, oversight of research and biosecurity education (Bezuidenhout 2012, p. 20; Selgelid & Rappert 2013, p. 277; Whitby et al. 2015, Fig. 7.2). In promoting a multifaceted web of measures to prevent the malicious use of novel innovations in biotechnology, the concept of the web of prevention quickly gained significant relevance in the academic debate on dual-use research and development. In this debate, the roles and responsibilities of research institutions and individual scientists within the web of prevention are stressed. As, for example, Miller (2018b) has argued, the notion of the web of prevention can be seen as an application of the concept of joint actions and collective moral responsibility and, in particular, of the concept of a layered structure of joint actions. All stakeholder groups within the web of prevention in question are jointly responsible for the collective end of preventing the production and use of biological weapons. Yet, the members of each stakeholder group perform a joint action(s) that is constitutive of the web, and each member of each group performs an individual action (and has a corresponding individual responsibility) in order to fulfil the collective end of the constitutive joint action in which they directly participate (Miller & Feltes 2018b, pp. 65–71).

However, the concept of a web of prevention has not been used exclusively in this specific context. Security researchers outside of the dual-use debate have referred to this concept and stressed the importance of an extensive set of stakeholders taking a multifaceted web of countermeasures against terrorist threats. For example, James Revill (2016) proposes a 'web of IED prevention' to combat the threat of terrorist attacks with improvised explosive devices (p. 93). By parity of reasoning, Feltes has deployed the concept of the web of prevention to analyse and improve the measures against the terrorist use of ricin, phosphine and americium (see Chapter 14 in this volume). Here, we advocate the deployment of the concept in relation to an entire CT strategy and, therefore, not restricted to terrorists' use of toxins.

This web of prevention against terrorist attacks requires the participation of at least the following institutions and constitutive institutional processes: government; security agencies, for example, the police, military and intelligence agencies; other state agencies, for example, finance departments; banks and other businesses; media, including social media companies; and citizenry. Each of these institutions has a collective end, or ends, and constitutive role structures to perform the necessary tasks to realize these ends; for example, governments develop an effective overall CT strategy, legislatures frame appropriate CT laws, intelligence agencies collect and analyse information, military forces engage in armed conflict against terrorist groups (if appropriate, as in the case of the Islamic State), police pursue criminal investigations, finance officers track money flows, journalists inform the public, and social media companies take down material inciting attacks.

In addition, countermeasures need to be designed and implemented for the purpose of responding to attacks if and when they occur, that is, on the assumption that prevention has failed. As Feltes (2020) has argued in relation to terrorist attacks using toxins, there are at least three groups of countermeasures: (1) measures to deny terrorists access to these substances, (2) measures to prevent the distribution of expertise that can be used to manufacture weapons with these substances, and (3) measures that are aimed at resilience and recovery in the aftermath of an attack with these substances.

In relation to the development of an effective overall CT strategy, it is important to address the issue of the motives underlying the establishment of a specific terrorist group and its support base; for example, the felt grievances of the Palestinian people over territory and statehood has led to the establishment of, and support for, Hamas. If there are legitimate grievances, then a key element of an overall CT strategy should presumably be to address these grievances.

The development of an effective CT strategy may require not only designing and implementing countermeasures and the like for existing institutions and institutional actors, but also the redesign of institutions and institutional roles, for example, the establishment of an agency to coordinate the CT strategy across various agencies that now have to cooperate more closely. Moreover, some of these institutional changes are likely to include new laws to restrict terrorist operations – for example, in relation to terrorist propaganda – and the granting of new legal powers to be attached to institutional actors – for example, new powers of detention for police. Accordingly, important questions may arise in relation to the moral, as opposed to pragmatic, justification for these new laws and legal powers because some of these additional laws and legal powers may compromise or curtail fundamental moral rights that are constitutive of liberal democracies. If so, elements of a liberal democracy's CT strategy may risk undermining the very society that it has been put in place to protect. In short, the collective moral responsibility to protect the members of the liberal democratic society from terrorist attacks needs to be discharged by institutional arrangements, including laws, that also respect individual moral rights.

NOTES

1. This is the collective end theory of joint action developed by Seumas Miller (1992, pp. 275–97, 2001, Ch. 2). For a related account, see Bratman (2014).
2. This view is developed and defended in Miller (2006, pp. 176–93).
3. This assumes that the intention is under their control and the intention causes the action in the right way. We cannot pursue the conceptual details of free and responsible action here, but see Paul et al. (1999) and Fischer (1986) for discussions of these issues.

4. Again, this assumes that the intention is under their control and the intention causes the action in the right way. See note 3.
5. See also Ludwig (2017).
6. See, for instance, Gross (2015).

REFERENCES

Bezuidenhout, Louise (2012),'Research Infrastructures, Policies and the "Web of Prevention": The Ethical Implications of Inadequate Research Environments', *Medicine, Conflict and Survival*, 28(1), 19–30.

Bratman, Michael (2014), *Shared Agency*, New York: Oxford University Press.

Feltes, Jonas (2020), *CBRN Threats, Counter-Terrorism and Collective Responsibility* (Unpublished PhD thesis), Delft University of Technology, the Netherlands.

Fischer, John M. (1986), *Moral Responsibility*, Ithaca: Cornell University Press.

Gross, Michael (2015), *The Ethics of Insurgency*, Cambridge: Cambridge University Press.

Ludwig, Kirk (2017), *From Plural to Institutional Agency*, New York: Oxford University Press.

Miller, Seumas (1992), 'Joint Action', *Philosophical Papers*, 21(3), 275–97.

Miller, Seumas (2001), *Social Action: A Teleological Account*, New York: Cambridge University Press.

Miller, Seumas (2006), 'Collective Moral Responsibility: An Individualist Account', in Peter A. French and Howard K. Wettstein (eds.), *Shared Intentions and Collective Responsibility*, vol. XXX in Midwest Studies in Philosophy, Boston: Wiley-Blackwell.

Miller, Seumas (2010), *The Moral Foundations of Social Institutions: A Philosophical Study*, New York: Cambridge University Press.

Miller, Seumas (2014), 'Police Detectives, Criminal Investigations and Collective Moral Responsibility', *Criminal Justice Ethics*, 33(1), 21–39.

Miller, Seumas (2018a), 'Joint Epistemic Action: Some Applications', *Journal of Applied Philosophy*, 35(2), 300–318.

Miller, Seumas (2018b), *Dual Use Science and Technology, Ethics and Weapons of Mass Destruction*, Cham: Springer.

Paul, Ellen F., Fred D. Miller, and Jeffrey Paul (eds.) (1999), *Responsibility*, Cambridge: Cambridge University Press.

Rappert, Brian, and Caitriona McLeish (2012), *A Web of Prevention: Biological Weapons, Life Sciences and the Governance of Research*, New York: Routledge.

Revill, James (2016), *Improvised Explosive Devices: The Paradigmatic Weapon of New Wars*, Cham: Springer.

Selgelid, Michael J., and Brian Rappert (eds.) (2013), *On the Dual Uses of Science and Ethics: Principles, Practices, and Prospects*, Canberra: ANU E Press.

Spencer, Richard (2014), 'Isil Carried Out Massacres and Mass Sexual Enslavement of Yazidis, UN Confirms', *The Telegraph*, 14 October 2014, accessed 18 May 2019 at https://www.telegraph.co.uk/news/worldnews/islamic-state/11160906/Isil-carried-out-massacres-and-mass-sexual-enslavement-of-Yazidis-UN-confirms.html.

Whitby, Simon M., Tatyana Novossiolova, Gerald Walther, and Malcolm R. Dando (2015), *Preventing Biological Threats: What You Can Do. A Guide to Biological Security Issues and How to Address Them*, University of Bradford, Bradford Disarmament Research Centre.

4. Kill, wound or capture: ethics considerations for counter-terrorism operations

Michael Robillard

1. INTRODUCTION

The twenty-first-century battle space continues to evolve at an ever-quickening pace. This claim holds true not only for conventional state-on-state conflicts but also for asymmetric conflicts. Part and parcel of this evolution is a set of new and emerging ethical problems, challenges and complexities. This is particularly true within counter-terrorism operations. While the domain of counter-terrorism operations involves a wide variety of distinct and important ethical issues – those related to torture and advanced interrogation, prisoner detention, intelligence gathering, collateral damage, national sovereignty and so on – this chapter focuses on some of the major ethical factors specifically related to kill, wound or capture criteria for high-value targets (HVTs).

Before we delve into this topic further, I must first make the disclaimer that I am not attempting to offer a specific recipe or rote set of ethical rules for counter-terrorism operators to strictly follow. Indeed, from Mill to McDowell to present debates regarding autonomous weapons, philosophers have often regarded the domain of ethics and morality to be, in principle, resistant to such simple reductions. If morality is to be anything at all, it must certainly be more than just a static set of inert algorithms and decision procedures.[1]

In Section 2, I briefly review core concepts in Just War Theory (JWT). In Section 3, I articulate some complexities that the phenomenon of terrorism creates for standard just-war thinking and explore several philosophical responses to such complexities. In Section 4, I examine some of the strengths and weaknesses of current ethical and legal guidelines specifically relating to kill, wound or capture criteria for HVTs in counter-terrorism operations. And in Section 5, I offer my own expansion of normative considerations as they relate to these various criteria and guidelines. From this analysis, my hope is that commanders and operators within the space of counter-terrorism will be

able to think about kill, wound or capture operations through a more nuanced and comprehensive lens of normative considerations and factors.

2. JUST WAR THEORY

Let us begin by laying out some foundational normative concepts typically related to the ethics of war. This is important so that we can isolate and articulate distinct and independent ethical values on the moral ledger relating to war as well as see how these ethical and normative reasons trade off against one another. Once this is done, we can then consider how these normative considerations trade off against other non-normative considerations, particularly those of a tactical or strategic nature.

To begin, let us first consider JWT or the just-war convention. JWT, broadly construed, can be understood as a philosophical position or set of positions existing between the two poles of pacificism and total war. Put simply, JWT rejects the idea that wars are always impermissible or, conversely, that they are always permissible and lacking any moral constraints whatsoever. JWT therefore asks the question: Given that people predictably will not act ideally, how ought we act?

JWT also posits theories about the special moral status of the context or domain of 'war'. Some philosophers, such as Michael Walzer, for instance, hold an exceptionalist view about war. In other words, Walzer believes that the moral rules that govern behaviour in a domestic setting with effective policing do not apply to the special domain of war, and that we instead must see war as a morally exceptional domain with a different set of moral rules and norms (Walzer 1977). Other philosophers, including Jeff McMahan, argue that, on the level of 'deep morality', the same morality that governs behaviour in domestic contexts holds for the context of war (McMahan 2009). This question of exceptionalism versus non-exceptionalism will have direct bearing on our assessment of terrorism.

2.1 *Jus ad Bellum*

JWT frequently makes the distinction between *jus ad bellum* and *jus in bello*.[2] *Jus ad bellum* concerns itself with the set of moral reasons and justifications for a group, political entity or nation state to justifiably go to war. *Jus in bello* concerns ethical behaviour on the battlefield once engaged in war. Arguably, the twenty-first-century informational age and the rise of global terrorism complicate our understanding and conceptualization of both of these spheres because terrorist organizations do not rise to the level of legitimately recognized nation states, nor do members of terrorist organizations typically adhere to standard *in bello* restraints as articulated by international law.

Typically, the sphere of *jus ad bellum* concerns the following major principles: just cause, right intention, legitimate authority, necessity, proportionality and likelihood of success.[3] According to most just-war accounts, as well as international law, satisfaction of all of these criteria is necessary and sufficient to fight a just war. Most pertinent to the issue of terrorism are the features of just cause and legitimate authority. Indeed, terrorist organizations often muddy the conceptual waters of both of these areas. Given that the organization in question is a terrorist organization, it almost by definition fails to meet the conditions of just cause as well as legitimate authority. However, by holding terrorist organizations and groups accountable to *ad bellum* just-war standards, the international community ends up tacitly acknowledging such groups as legitimate state actors. Such a paradox therefore motivates the argument to legally and morally treat terrorist groups as either criminals under the conceptual rubric of domestic policing, or under some alternative category other than the standard law of armed conflict (LOAC) conventions, which pertains to legitimate combatants. To be clear here, this is not to conflate the first-order ethical principles of *jus ad bellum* with international legal norms and conventions governing armed conflict between nation states. Rather, the point here is simply to highlight the paradox that emerges when attempting to appeal to these moral principles to explain the wickedness of terrorist actors, while simultaneously attempting to deny such actors recognition of legitimacy.

2.2 *Jus in Bello*

As stated previously, traditional JWT makes the moral distinction between the ethics of a state going to war (*jus ad bellum*) and the ethics of soldier behaviour in war (*jus in bello*). Fundamental to the conceptual space of *jus in bello* is the notion of the 'moral equality of combatants' (MEC). The concept of the MEC is that, independent of the justness of the war itself, soldiers on both sides of a military conflict can nonetheless 'fight well' if they exercise ethical restraint with respect to using only necessary and proportionate force towards their designated enemy, discriminate between combatants and non-combatants, exercise caution to minimize collateral damage, and respect the basic rights of prisoners of war (see Frowe and Lazar 2018). These *in bello* moral reasons are often codified in standard soldier rules of engagement and in international LOAC, such as the Geneva Conventions.

Terrorist organizations obviously muddy the conceptual waters of these standard *in bello* distinctions, as well, since terrorists do not fight on behalf of a legitimate nation state, do not wear uniforms designating themselves as combatants, and often operate in battle spaces that overlap and intertwine with the civilian populace. These are arguably necessary for, or implied in, the *jus in bello* thinking, which places pressure on how to apply these principles to

a context of terrorism. What is more, a given individual's causal contribution to terrorist organizations will often be scalar and networked in nature and fail to admit to the binary combatant/non-combatant distinction on a battlefield as witnessed in twentieth-century conventional warfare. These are all ethical nuances of the twenty-first-century battle space of which commanders must be cognizant.

3. TERRORISM AND THE JUST-WAR CONVENTION

Given that terrorist organizations do not fit cleanly within the standard moral and legal frameworks of either domestic policing or the LOAC, it is unclear how exactly we should ethically and efficaciously respond to such newly arising phenomenon within the global theatre. Should we treat terrorist organizations like they are domestic organized crime syndicates and therefore subject to moral conventions of domestic policing? Should we instead treat them like they are standard legal combatants on a battlefield and therefore subject to international legal norms and guidelines pertaining to legal and legitimate combatants? Should we treat them like they are some sort of entity with a moral status existing somewhere in between these two poles? Or, should we treat them as being fully outside both the normative and contractual space of domestic policing and the just-war convention altogether? Several prescriptions and responses to these questions have been offered that range in their degree of severity, permissiveness and stringency.

Isaac Taylor, for instance, in his book, *The Ethics of Counterterrorism*, argues that what is fundamentally needed is an exceptionalist approach to counter-terrorism, which requires a suspension of traditional just-war rules and frameworks. He writes,

> A distinct set of rules are needed to govern counterterrorist efforts by governments and other actors. Thought of in this way, counterterrorism is subject to what Fritz Allhoff calls an 'ethics of exceptionalism' according to which the War on Terror, through its novel face and extreme stakes suggests…that we need to make exceptions to traditional norms. (Taylor 2018)

Borrowing from Gerard A. Cohen's distinction between 'fundamental (normative) principles' and 'rules of regulation', Taylor argues for exceptionalism with respect to the latter.[4] For Cohen, first-order normative reasons concern 'deep morality', while rules of regulation concern the second-order norms we collectively use to prudentially attend to and realize these first-order values. In this sense, Taylor still acknowledges the normative reality of deep first-order deontological, consequentialist and virtue-theoretic reasons in war but calls

for a suspension of the standard just-war norms and conventions typically applied to conventional warfighting and uniformed soldiers when it comes to counter-terrorism operations. In this sense, Taylor is an exceptionalist about 'rules of regulation' but not an exceptionalist about fundamental (normative) principles in war, as Walzer is. Taylor is not specific about which norms and conventions to suspend or revise, but he, like Allhoff, at least suggests that the moral severity of global terrorist threats would justify and warrant such a suspension.[5]

Other just-war theorists, such as Daniel Statman, argue for a similar suspension of standard *in bello* norms when it comes to conducting counter-terrorism operations. In Statman's view, the just-war convention is grounded in a set of explicit and/or tacit contracts between legitimately recognized nation states and therefore, by proxy, between these states' legitimate combatants. Likewise, terrorist actors and organizations are simply outside of the war convention altogether and therefore not deserving of such moral restraint in terms of targeting, prisoner detention and so on (Statman 2004).

4. PRESENT PRESCRIPTIONS

Despite these 'exceptionalist' views with respect to *in bello* treatment of terrorist groups and actors, robust national and international legal norms have been nonetheless devised to regulate targeting of such agents. Whereas past twentieth-century legal norms constraining soldier behaviour *in bello* were largely indexed to a presumed context of conventional militaries fighting on a designated battlefield, and with necessary and proportionate violence largely being exchanged exclusively between uniformed soldiers, the increasingly unconventional character of twenty-first-century counter-terrorism operations has blurred and obfuscated such norms. Indeed, with terrorist networks operating within and amongst large urban and civilian contexts, and without official uniforms or a designated battlefield, the epistemic threshold for what constitutes liability for justified *in bello* targeting has largely been explained in terms of the 'imminent threat' in which such terrorist groups' collective behaviours predictably might eventuate. Furthermore, targeting criteria in international norms have been further explained in terms of the directness of knowledgeable and/or culpable participation within a given terrorist network.

Let us now therefore turn to look at some of the specific legal and normative guidance and language with respect to the killing, wounding or capturing of HVTs.

4.1 2001 Authorization for Use of Military Force

The legal justification for US counter-terrorism efforts in recognized 'areas of active hostilities' finds its grounding in the 2001 AUMF (Authorization for Use of Military Force against Terrorists). The 2001 AUMF granted the President of the United States legal authority to use all necessary and appropriate force against those he determined 'planned, authorized, committed, or aided' the terrorist attacks on 11 September 2001 (AUMF 2001). This legal document, along with its 2002 Iraq-specific instantiation, has provided the main legal justification for continued US counter-terrorism efforts for the last four administrations.

4.2 International Committee of the Red Cross Standards

In addition to the 2001 AUMF, the International Committee of the Red Cross (ICRC) provides another legal standard for justified use of force in counter-terrorism operations. The ICRC describes the following activities as meeting the threshold of harm connected with direct participation in hostilities:

> Capturing, wounding or killing military personnel; damaging military objects; or restricting or disturbing military deployment, logistics and communication, for example through sabotage, erecting road blocks or interrupting the power supply of radar stations. Interfering electronically with military computer networks (computer network attacks) and transmitting tactical targeting intelligence for a specific attack are also examples. The use of time-delayed weapons, such as mines or booby traps and remote-controlled weapon systems, such as unmanned aircraft, also 'directly' causes harm to the enemy and, therefore, amounts to direct participation in hostilities. (ICRC 1989)

The ICRC also provides that some persons who regularly engage in what it calls a 'continuous combat function' may be considered members of non-state-organized armed groups who can be targeted as combatants at any time. Such persons are those who are continuously engaged in 'the preparation, execution, or command of acts or operations amounting to direct participation in hostilities'. The ICRC explains:

> a continuous combat function may be openly expressed through the carrying of uniforms, distinctive signs, or certain weapons. Yet it may also be identified on the basis of conclusive behaviour, for example where a person has repeatedly directly participated in hostilities in support of an organized armed group in circumstances indicating that such conduct constitutes a continuous function rather than a spontaneous, sporadic, or temporary role assumed for the duration of a particular operation. (ICRC 1989)

4.3 US Department of Defense Standards

Much of the behaviour that the United States regards as the basis for targeting is more expansive than the ICRC standard. The US Department of Defense (DoD) *Law of War Manual* states:

> Being part of a non-State armed group that is engaged in hostilities against a State is a form of engaging in hostilities that makes private persons liable to treatment in one or more respects as unprivileged belligerents by that State. Being part of a non-State armed group may involve formally joining the group or simply participating sufficiently in its activities to be deemed part of it. (US DoD 2015)

In such cases, the United States has identified someone who is engaged in behaviour that appears to be directed to the eventual infliction of harm by themself or others. The extent that the US regards such a person as posing a threat that is imminent is based on a variety of factors. These factors include:

> the nature and immediacy of the threat; the probability of an attack; whether the anticipated attack is part of a concerted pattern of continuing armed activity; the likely scale of the attack and the injury, loss, or damage likely to result therefrom in the absence of mitigating action; and the likelihood that there will be other opportunities to undertake effective action in self-defense that may be expected to cause less serious collateral injury, loss, or damage. (US DoD 2015)

Given this expansive definition of imminence, waiting to act until obtaining information about identity may well not impose any military cost at all.

4.4 2013 Presidential Policy Guidance

A more updated justification for US counter-terrorism efforts as well as an attempt to account for new drone technologies was articulated in the Obama administration's Presidential Policy Guidance (PPG) of 2013. The PPG of 2013 states that: 'The standard operating procedures for when the United States takes direct action, which refers to lethal and non-lethal uses of force, include capture operations against terrorist targets outside the United States and areas of active hostilities'.

The 2013 PPG further outlines certain guidelines for permissible lethal targeting as follows:

> In addition to the several requirements previously announced for all uses of force outside Afghanistan, Iraq and Syria (i.e., that the target poses 'a continuing, imminent threat to U.S. persons'; near certainty that the target is present; near certainty that non-combatants will not be injured or killed; an assessment that capture is not feasible at the time of the operation; an assessment that the relevant governmental

authorities in the country where action is contemplated cannot or will not effectively address the threat to U.S. persons; and an assessment that no other reasonable alternatives exist to effectively address the threat to U.S. persons).

4.5 Trump Administration Standards

Upon taking office, the Trump administration began loosening some of the constraints on lethal targeting established by the 2013 PPG. Such loosening included moving control of some targeted killing decision-making from Title 10 DoD jurisdiction to Title 50 Central Intelligence Agency jurisdiction, arguably a move towards greater opacity. This theme of greater opacity has also been reflected in the Trump administration's decision to depart from the Obama-era norm of reporting targeted killing effects to the general public (Dilanian and Kube 2019). Lastly, the Trump administration's dropping of the 'continuing, imminent threat' standard for targets and 'near-certainty' standard for non-combatant collateral damage has generated much concern from many critics (Serle 2017).

While none of these standards should be seen as finalized, they at least give some substantive legal and moral precedent to which command decisions in counter-terrorism operations can be indexed.

5. ADDITIONAL ETHICAL CONSIDERATIONS

Now that we have investigated the major conceptual moving parts of JWT, have explored various ethical arguments relating specifically to terrorism, and have examined existing legal and moral frameworks for kill, wound or capture operations, in this final section, I discuss several under-acknowledged normative factors relevant to counter-terrorism operations. While this set of moral considerations is not meant to be exhaustive, I offer these for commanders to think about in conjunction with the ethical and legal frameworks already examined so as to augment their in-theatre planning and decision-making.

5.1 Feasible Alternatives

Perhaps the most ethically relevant and ethically important factor in relation to kill, wound or capture operations is the consideration of feasible alternatives. Indeed, were it the case that the only other way to kill, wound or capture an HVT would be to commit a large-scale conventional force to a bloody, costly and drawn-out 'boots on the ground' type of engagement, then, all things considered, it seems like it would be morally preferable to resort to some sort of targeted killing option instead. Such considerations, of course, will be sensitive to the specifics of a given mission and the moral constraints of *in bello* propor-

tionality and necessity, and will have to be a judgement call for commanders to make on an ongoing basis given tactical and strategic constraints. Such considerations will be further sensitive to the macro *ad bellum* conditions of last resort and proportionality throughout the entirety of the campaign or conflict.

5.2 Personality versus Signature-Strike Thresholds

A second ethical consideration that is similarly pertinent to kill, wound or capture operations is the question of what level of epistemic threshold is sufficient to justify the use of targeted killing. Two standards, that of personality strikes versus signature strikes, seem to be in question.[6] Personality strike criteria would require metadata on pattern-of-life behaviour plus positive identification of the particular person before targeting. Alternatively, signature-strike criteria would set the standard of justifiable targetability at the level of metadata 'pattern-of-life' behaviour that suggested imminent threat but not, as it were, at the level of positive personal identification. Hence, these two standards would generate two epistemic thresholds for justifiable targeting and would also generate two sets of possibilities for false negatives and false positives.

5.3 Morality Hacking

In his article 'Moral Coercion', Saba Bazargan-Forward (2014) presents several hypothetical cases whereby an unjust enemy uses standard just-war ethical constraints as a means to trap his just opponent in a tactically disadvantageous situation. The use of human shields in battle represents perhaps the clearest example of such an instance of war's morality being 'hacked'. Given such a case, one might then wonder if, after some number of instances, it would be ethically permissible for a military leader to loosen normal *in bello* constraints to counter such a dirty tactic. Such a loosening of *in bello* constraints seems, at least in principle, plausible at some proportionality threshold if a significantly weighty moral good was at stake, such as the possibility of losing the entire war.[7] Such loosening seems further justifiable if the exceptionalist view of scholars such as Taylor (2018) and Statman (2004) is, in fact, correct.

5.4 Respect for Soldiers

Another morally relevant factor connected to kill, wound or capture operations is that of respect for soldiers. Indeed, in managing the moral costs and benefits of certain operational actions, we must not forget the moral stakes connected to the risk of soldier lives. While it is the sworn job of soldiers to risk their lives

in service of the higher mission, it is important to note that the taking on of the moral status of combatant does not give the military institution carte blanche permission to send soldiers 'into the guns' needlessly or recklessly to achieve a given kill, wound or capture objective. Such missions will obviously depend on feasibility as well as a soldier's or unit's willingness to take on additional moral risks. That being said, despite such willingness, regard for soldiers' objective well-being should still count somewhere within any commander's ethical and operational calculus.[8]

5.5 Narrative Dominance and the Communication of Reasons

Another morally relevant consideration for kill, wound or capture operations is that of narrative dominance. Put simply, in conventional as well as asymmetric conflicts, the use of violence in warfare carries ethical weight and pragmatic utility not only in virtue of its eliminative capacities (that is, killing the enemy and smashing their equipment), but also with respect to its communication of reasons (Underwood 2019). Indeed, from a terrorist actor's perspective, the effectiveness of a violent terrorist act will largely hinge on the act's ability to effectively communicate reasons and narratives to the populace of whom they are seeking to gain control. This means that narrative dominance is utterly essential in counter-terrorism campaigns. Accordingly, the managed use of violence in counter-terrorism operations can carry with it both moral and pragmatic import with respect to the narrative or symbolic meaning that is communicated. Specifically, the choice to kill versus wound versus capture carries morally important signifying value. Indeed, a spectacular show of violent force can possibly be justifiable insofar as it carries significant future deterrence value, thereby preventing future protracted conflict. Similarly, handling an HVT target 'with kid gloves', could possibly hinder counter-terrorism operations by emboldening terrorists to act out with no fear of repercussion. Conversely, an overly aggressive killing of an HVT could create the long-term symbolic effect of radicalizing the area and trading tactical wins for strategic setbacks. Some criticisms of the Obama administration's over-reliance on drone targeted killings suggests such blowback and radicalization dangers (Pilkington and MacAskill 2015).

Lastly, commanders must consider what narrative or signifying message is being sent to civilians on the battlefield, as well as to allied partners and to the world at large. Indeed, given the increased presence of smartphones, social media and the so-called 'velocity of information', the narrative and symbolic effects of kill, wound or capture missions will be predictably felt beyond the contained epistemic space of the 'in-theatre' battle space.

5.6 Citizenship

Another moral factor relevant to kill, wound or capture missions is that of the citizenship of the target. Perhaps the clearest example of citizenship complicating the morality of targeted killing decision-making was that of the 2011 drone strike on al-Qaeda member Anwar al-Awlaki in Yemen. What arguably made the strike on Awlaki more morally complicated than regular drone strikes on terrorist HVTs was the fact that Awlaki was also an American citizen. Given this, the targeting of Awlaki, in some views, marked the first unprecedented instance of the extrajudicial killing of an American citizen without due process and technically outside of an area of active hostilities. One might wonder what kind of precedent-setting and signalling effects such targeting decisions have on the world community as well as allied partners.[9]

5.7 National Sovereignty

Another morally relevant issue connected to kill, wound or capture operations is that of national sovereignty. Unlike in conventional wars, where enemy combatants can often be legally and morally targeted on a recognized battlefield, effective counter-terrorism campaigns often involve operations that occur within the physical space of other legitimately recognized nation states. Without an official war occurring, such operations therefore run the risk of infringing other nation states' sovereignty rights.[10] Hence, if a legitimately recognized nation state seems unwilling or unable to effectively kill, wound or capture terrorist groups or actors operating within their jurisdiction, then a decision must be made that considers the moral and strategic trade-offs between temporary sovereignty infringement versus the good to be achieved by thwarting an imminent attack.

The 2011 raid on Osama bin Laden in northern Pakistan and the 2011 Awlaki drone strike in Yemen perhaps stand as the two most prominent counter-terrorism operations to date where sovereignty rights were temporarily infringed in order to kill, wound or capture a significant HVT. In assessing the moral dimensions of such counter-terrorism acts, commanders must consider not only the costs and benefits of the given act in isolation, but also the signalling effects and precedent-setting that such acts have for the international community.

5.8 Complicity

Closely related to the issue of national sovereignty is the question of the moral relevance of complicity. This question of complicity applies not only to macro-level knowledge and tolerance of terrorist groups operating within

a nation state's internal borders, but also micro-level knowledge of civilians within local neighbourhoods.[11] Put simply, we might ask, 'what degree of knowledge, wilful ignorance and/or inaction with respect to terrorist actions must a nation or non-combatant demonstrate before sovereignty rights or individual bodily rights can be justifiably infringed?' At the macro level, a certain degree of complicity and failure to sufficiently attend to or aid in counter-terrorism operations in one's own jurisdiction could feasibly justify a temporary infringement of nation state sovereignty. At the micro level, failure to report known terrorist groups and/or to physically distance oneself from known terrorist areas could feasibly justify loss of life as a result of collateral damage.

5.9 Stipulating Conventions

One possible solution to the issue of complicity, at both the macro and micro levels, is the practice of stipulated conventions or stipulated norms. While it has been a much-debated issue within just-war dialogues as to the moral force of conventions, it seems at least plausible that stipulated conventions and norms could have moral force unto themselves after a sufficient number of iterations.[12] For instance, in a tactical scenario, a leaflet campaign could be used over a given neighbourhood alerting villagers to a terrorist cell operating in their area and warning them that they must exit the area immediately, otherwise they will be seen as co-conspirators. While it is still debatable, some ethicists will argue that such stipulated warnings then place a new moral constraint on complicit non-combatants, obligating them to act to get out of the region in question. Such a practice could then be possibly used by counter-terrorism operators when dealing with the issue of a complicit populace.

5.10 Sunk Costs

A final moral consideration pertaining to kill, wound or capture operations is that of sunk costs. This issue applies at both the macro and micro levels. On the macro level, commanders must assess the 20 years of continuous combat operations in the Middle East and whether or not such *ad bellum* moral justifications have an expiration date or at least ought to be discounted as time goes on. Other philosophers argue that no such discounting should be made and that new commanders should roll sunk costs into their justifications for future combat operations. Still other philosophers contend that sunk costs ought to be completely dismissed when commanders take command, and the moral calculus ought to be reset.[13] This same question of whether not to discount, roll forward or reset repeats itself at the *in bello* level whenever commanders take over from their operational predecessors. Accordingly, commanders for

counter-terrorism operations must take such *ad bellum* and *in bello* sunk considerations into account.

While these moral considerations are not meant to be fully comprehensive, they at least function to supplement existing legal and moral guidance within the space of kill, wound or capture operations.

6. CONCLUSION

In this chapter, we have reviewed some of the major moral and pragmatic elements of kill, wound or capture criteria for counter-terrorism operations. Specifically, we have investigated JWT, broadly construed, particularly the facets of *jus ad bellum* and *jus in bello*, and we have explored various philosophical responses to terrorism in particular. Furthermore, we have investigated existing legal guidance on targeted killing for the US and internationally. Lastly, we have looked at a set of other under-acknowledged normative factors pertinent to kill, wound or capture operations. While these moral considerations are not intended to be exhaustive when it comes to decision-making for kill, wound or capture operations, I nonetheless offer these moral considerations to augment existing counter-terrorism thinking and planning since they are not often explicitly expressed in present targeted killing guidelines. I therefore leave it to present and future counter-terrorism commanders and operators to decide what level of stringency makes the most sense to them, given other competing ethical and strategic priorities.

NOTES

1. What I am attempting to do in this chapter, then, is to clearly articulate and make explicit the distinct values that are often in play during kill or capture operations to explore how these various moral values trade off against one another, and to explore how these various normative values relate to other non-normative considerations: epistemic, pragmatic, and otherwise.
2. In recent years, philosophers including David Rodin have turned their attention to *jus ex bello* (the ethics of exiting a war justly). Other philosophers, such as Larry May, have focused on issues of *jus post bellum* (justice after war). See Rodin (2015) and May (2012).
3. For a general overview of the contemporary just-war landscape, see Frowe and Lazar (2018).
4. Taylor's reference to Cohen comes from Cohen (2008).
5. Taylor's reference to Fritz Allhoff comes from Allhoff (2012).
6. Whereas a 'personality' strike involves the epistemic criteria of knowing the specific person, usually by name, who is being lethally targeted, 'signature strikes' set the epistemic bar much lower, relying on less-than-positive identification of a specific target, such as metadata analysis and 'pattern of life' criteria. For an analysis of some of the moral worries surrounding signature strikes, see Rohde (2015).

7. For a good treatment of such threshold deontology issues, see Jesper Ryberg (2010).
8. The degree of voluntariness of soldier commitment is, of course, questionable these days. See Robillard and Strawser (2016).
9. For an in-depth treatment of the legal and moral issues surrounding the Awlaki killing, see Chesney (2010).
10. For a good philosophical treatment of such sovereignty concerns with respect to counter-terrorism operations, see Strawser (2014).
11. For treatments of questions of complicity, culpability and liability, see Bazargan-Forward (2017).
12. For debates on the moral force of conventions, see Lewis (1969) and Verbeek (2008).
13. For a thorough treatment of these various views on sunk costs, see Tadros (2018).

REFERENCES

Allhoff, Fritz (2012), *Terrorism, Ticking Time-Bombs, and Torture: A Philosophical Analysis*, Chicago: University of Chicago Press.

Authorization for Use of Military Force (2001, September 18), Government Printing Office, accessed 15 April 2020 at http://www.gpo.gov/fdsys/pkg/PLAW-107publ40/pdf/PLAW-107publ40.pdf.

Bazargan-Forward, Saba (2014), 'Moral Coercion', *Philosopher's Imprint*, 14(11), 1–18.

Bazargan-Forward, Saba (2017), 'Complicity', in Marija Jankovic and Kirk Ludwig (eds), *The Routledge Handbook of Collective Intentionality*, Abingdon: Routledge, 327–37.

Chesney, Robert (2010), 'Who May Be Killed; Anwar Awlaki as a Case Study in the International Legal Regulation of Lethal Force', *Yearbook of International Humanitarian Law*, 13, 3–61, https://doi.org/10.1007/978-90-6704-811-8_1.

Cohen, Gerard A. (2008), *Rescuing Justice and Equality*, Cambridge, MA: Harvard University Press.

Dilanian, Ken, and Courtney Kube (2019, 6 March), 'Trump Cancels Obama Policy of Reporting Drone Strike Deaths', NBC News, accessed 15 April 2020 at https://www.nbcnews.com/politics/donald-trump/trump-cancels-obama-olicy-reporting-drone-strike-deaths-n980156.

Frowe, Helen, and Seth Lazar (2018), *The Oxford Handbook of The Ethics of War*, New York: Oxford University Press.

International Committee of the Red Cross (1989), Article 57, Additional Protocol I to the Geneva Conventions Section (2)(a)(i), accessed 15 April 2020 at https://www.icrc.org/en/doc/assets/files/other/icrc_002_0321.pdf.

Lewis, David (1969), *Convention*, Cambridge, MA: Harvard University Press.

May, Larry (2012), 'Remarks by Larry May', *Proceedings of the Annual Meeting of the American Society of International Law*, 106, 332–34, https://doi.org/10.5305/procannmeetasil.106.0332.

McMahan, Jeff (2009), *Killing in War*, Oxford: Clarendon Press.

Pilkington, Ed, and Ewan MacAskill (2015, 18 November), 'Obama's Drone War a "Recruitment Tool" by ISIS, Say US Air Force Whistleblowers', *The Guardian*, accessed 26 March 2021 at https://www.theguardian.com/world/2015/nov/18/obama-drone-war-isis-recruitment-tool-air-force-whistleblowers.

Presidential Policy Guidance (2013), *Procedures for Approving Direct Action against Terrorist Targets Located outside the United States and Areas of Active Hostilities*, accessed 15 April 2020 at https://www.aclu.org/sites/default/files/field_document/presidential_policy_guidance.pdf.
Robillard, Michael, and Bradley J. Strawser (2016), 'The Moral Exploitation of Soldiers', *Public Affairs Quarterly*, 30(2), 171–95.
Rodin, David (2015), 'The War Trap: Dilemmas of *jus terminatio*', *Ethics*, 125(3), 674–95.
Rohde, David (2015, 28 April), 'What the United States Owes Warren Weinstein', *The Atlantic*, accessed 26 March 2021 at https://www.theatlantic.com/international/archive/2015/04/warren-weinstein-drones/391655/.
Ryberg, Jesper (2010), 'Mass Atrocities, Retributivism, and the Threshold Challenge', *Res Publica*, 16(2), 169–79.
Serle, Jack (2017), 'Trump, Obama, and the Future of Targeted Killing', *The Bureau of Investigative Journalism*, accessed 15 April 2020 at https://www.thebureauinvestigates.com/stories/2017-01-19/trump-obama-and-the-future-of-targeted-killing.
Statman, Daniel (2004), 'Targeted Killing', *Theoretical Inquiries in Law*, 5(1), 179–98.
Strawser, Bradley J. (2014), *Killing Bin Laden: A Moral Analysis*, New York: Palgrave Macmillan.
Tadros, Victor (2018), 'Past Killings and Proportionality in War', *Philosophy & Public Affairs*, 46(1), 9–35.
Taylor, Isaac (2018), *The Ethics of Counterterrorism*, New York: Routledge.
Underwood, Robert (2019, 12 March), 'Can Soldiers Justify Killing Some as a Means to Influence the Decisions of Others?', *Practical Ethics*, accessed 26 March 2021 at http://www.bioethics.net/2019/03/oxford-uehiro-prize-in-practical-ethics-question-can-soldiers-justify-killing-some-as-a-means-to-influence-the-decisions-of-others/.
US Department of Defense (2015), *Law of War Manual*, accessed 15 April 2020 at https://dod.defense.gov/Portals/1/Documents/pubs/DoD%20Law%20of%20War%20Manual%20-%20June%202015%20Updated%20Dec%202016.pdf?ver=2016-12-13-172036-190.
Verbeek, Bruno (2008), 'Conventions and Moral Norms: The Legacy of Lewis', *Topoi*, 27(1–2), 73–86.
Walzer, Michael (1977), *Just and Unjust Wars*, New York: Basic Books.

5. Accountability for targeted killing

Mary B. DeRosa and Mitt Regan

1. INTRODUCTION

For almost 20 years, the United States (US) has engaged in the world's most extensive program of targeted killing strikes outside of active battlefields. These operations have generated extensive legal and ethical debate that raises crucially important issues. However, we set these to one side in this chapter to focus on how to ensure that government can be held accountable when it engages in such operations.

In May 2013, the Obama administration announced a Presidential Policy Guidance (PPG) with standards for targeted strikes outside areas of active hostilities. These included the requirement that a target present a "continuing imminent threat to US persons," that capture of a target be infeasible, and that there be near certainty of identification of a target and near certainty that there will be no civilian casualties (White House 2013).

The Trump administration did not disclose its targeting criteria, but reports indicate that it relaxed the definition of a permissible target, as well as the requirement that the target be identified with near certainty. The requirements that capture be infeasible and near certainty of no civilian casualties appear to have continued in place during the Trump administration (Savage & Schmitt 2017; Hartig & Tankel 2019). In January 2021, the new Biden administration suspended the Trump administration guidance and began a review of targeting policy.

The PPG represents an effort to ensure that targeted killing occurs subject to publicly announced criteria, consistent with basic requirements of the rule of law. The question remains, however: how best can the US be held accountable for adherence to these standards? Targeted strikes are conducted based on intelligence that government may not be able to fully disclose, even to all members of Congress. Some critics thus contend that the targeted killing program remains shrouded in secrecy that prevents meaningful accountability.

This chapter discusses the debate over accountability for the US targeted killing program, and the potential mechanisms for enhancing it. We explore accountability, first, with respect to the decision to designate an individual

as a target and, then, with the strike operation itself. Our focus is on the US program, but our aim is to illuminate the complex considerations involved in attempting to strengthen accountability of any government that conducts these types of operations.

2. TARGET-DESIGNATION PHASE

2.1 Standards and Process for Target Designation

The target-designation phase of a targeted killing operation takes place largely in Washington, DC. It involves decisions about whether a potential target satisfies the policy criteria to be designated for a lethal use of force; development of an operational plan, including assessment of whether it is possible to satisfy operational requirements; legal review; final policy-level approval; and some external oversight (White House 2013).

The first step is to determine whether a potential target satisfies relevant threat criteria. As a threshold matter, the target must be a member of an organization for which there is legal authority to use force under international and domestic law. The Obama administration focused lethal targeting of the sort we address here against members of al-Qaeda and its affiliates or associated forces (White House 2016b, pp. 5–7). There are indications that the Trump administration adopted a broader view of the organizations against which this tool would be employed. For example, in January 2020, the Trump administration conducted a successful lethal strike against Iranian Major General Qasem Soleimani while he was in Iraq (White House 2020). Although the US described General Soleimani as a terrorist, he was not affiliated with al-Qaeda.

To be targeted, the Obama administration required an individual to pose a "continuing, imminent threat to US persons." The PPG does not define this standard, but the Obama administration explained that not all terrorists overseas pose this threat (White House 2016b, p. 25). Generally, those deemed a continuing, imminent threat were senior operational leaders of their organization, those whose unique operational skills – such as bomb making – made them a particular threat, or others actively planning attacks against the US. As a policy matter, the Obama administration declined to target low-level "foot soldiers," except those in the process of preparing for or executing terrorist attacks (Savage & Schmitt 2017). The Trump administration reportedly eliminated this "continuing, imminent threat" requirement, but it is not clear what standard it adopted for determining whether an individual poses a sufficient threat to be designated for targeting by lethal force.

The PPG contains additional significant policy restrictions relevant to the operational phase of lethal targeting. It requires "near certainty" both that the targeted individual is present at the scene and that the strike will not injure or

kill non-combatants. The PPG also permits targeting only if capture is infeasible at the time of the operation and there are no reasonable alternatives to address the threat that the target poses. These requirements must be considered during operational planning, but operators in the field make the final assessment at the time of the operation. In addition, the PPG addresses the possibility that an operation will violate the sovereignty of the country in which the strike is planned. If a country has not consented to a strike in its territory, the planners may only move forward with an operation if relevant government authorities in the country "cannot or will not effectively address the threat to US persons" (White House 2013).

The restrictive PPG standards and processes apply only to proposed strikes that take place outside "areas of active hostilities." The PPG does not define what constitutes an area of active hostilities. Initially that term was used to refer only to the conflicts in Afghanistan and Iraq, but over time several additional locations were designated as such, and thereby removed from the PPG's more restrictive requirements.

The operational plan undergoes policy and legal review within the executive branch before final approval. The PPG requires each lethal targeting proposal to go through interagency legal review and several layers of National Security Council-led policy review, ending with a high-level "principals" review, led by the National Security Advisor and including cabinet-level national security officials. Each level includes senior defense and intelligence officials, as well as leaders from departments such as State, Justice, and Homeland Security, and, often, the US Mission to the United Nations, who are likely to bring a distinctive perspective to the discussion. If all principals agree that a proposed lethal targeting operation should proceed, the president is informed of the decision before the operation. If there is disagreement among principals, or if the target is a US person, the PPG requires the president to make the final decision.

The Trump administration reportedly eliminated much of this interagency review, creating a process that is more streamlined and efficient, but that benefits less from a range of perspectives from outside of the defense and intelligence communities. In addition, the Trump administration lowered the approval level for lethal targeting, no longer requiring presidential approval for most proposed operations. Instead, the decision about whether to move forward with many lethal targeting operations likely rested with military combatant commanders and, thus, was not subject to interagency review by senior-level officials (Savage & Schmitt 2017; Hartig & Tankel 2019).

Public reporting during the Obama administration of the process for selecting and approving targets for lethal operations was hesitant and sporadic but increased over time. By 2016, there had been significant reporting of the process by which these decisions were made – including release of the PPG – and of the legal analysis that accompanied them. To the degree there was

a trend toward openness in the Obama administration, it reversed with the Trump administration. After 2017, there was little public discussion of the process or standards for these operations (Atherton 2020).

External oversight of the target-designation process is minimal. The judicial branch plays virtually no role in overseeing decisions about lethal targeting operations. Although some litigants have attempted to engage the courts, particularly on the targeting of US citizens, the courts have shown little inclination to become involved, typically dismissing cases on jurisdictional grounds before reaching the substantive issues.[1]

The executive branch has had more interaction with the US Congress about these matters. The PPG requires notification to relevant congressional committees of new operational plans for lethal targeting or changes to those plans. Although there is no requirement of pre-strike notice to Congress, the operational agencies must inform the relevant committees after conducting an operation pursuant to approved plans. The PPG also requires updates to certain members of relevant congressional committees at least every three months about individuals approved for lethal targeting (White House 2016b). It is not clear whether the Trump administration continued to follow these notification practices.

2.2 Accountability Concerns

The process described above for selecting targets and developing an operational plan includes several substantive and procedural protections. Nonetheless, considering the gravity of these decisions, the process raises many accountability concerns.

How the executive branch sets standards for lethal targeting – as a matter of policy, with little oversight or transparency – leads to internal and public uncertainty about their substance and durability. This reduces the credibility of the resulting operations. As noted above, the Obama administration improved in this regard over its eight years, and the PPG represented a significant advance in terms of rigor and clarity about the standards and process for lethal targeting. The fact that the guidelines eventually were released publicly added to the policy's legitimacy.

However, criticisms of the Obama administration's approach remain. First, some of the PPG standards are ill-defined. One noteworthy example is the notion of "areas of active hostilities." The processes set forth in the PPG apply only outside of "areas of active hostilities," but the document does not define that term, and there appears to have been confusion even within the Obama administration about who makes that determination and on what basis. The strict requirements of the PPG do little good if the situations in which they apply are defined away. Toward the end of the Obama administration, it was

unclear, at least publicly, exactly how broadly the PPG standards applied, amid reports that Syria, Somalia, and parts of Libya were considered to be "areas of active hostilities."

Of even greater concern, unlike with a statutory or other legal requirement, a president can change policy at any time. Although there are political disincentives for a president to withdraw or change his own publicly announced policy, that is less true when the presidency changes hands. As noted, the Trump administration withdrew the PPG and never made its own policies public. Although little is clear about the Trump administration's process, we understand that it eliminated the requirement that a proposed target pose a "continuing, imminent threat" to the US and the need for near certainty about the target's location. It is not clear what replaced these standards. It also eliminated many steps and participants in the executive branch's decision-making process, and its policy on congressional notification is unclear (Savage & Schmitt 2017; Hartig & Tankel 2019). This lack of clarity about these policies made public accountability difficult.

Several aspects of the internal executive branch decision-making process raise concerns about the quality of those decisions and how likely they are to reflect broader policies and the rule of law. First, any process focused on approving individual targets and operational plans risks skewing the discussion by underplaying the aggregate strategic impact of these operations. Although each targeted strike raises tactical and policy issues, in the aggregate, the strikes raise broader strategic questions about adherence to the rule of law and about US credibility in the Middle East and beyond. When discussions focus only on individual strikes, these broader issues can get lost. This concern arises even with the extensive PPG procedures, but the PPG at least guaranteed the participation of officials with a variety of perspectives, including officials from the State Department and the US Mission to the United Nations, who are institutionally more attuned to these broader concerns and more likely to inject them into the discussion (White House 2016a).

The PPG process has been criticized, with some justification, for being overly bureaucratic (Savage & Schmitt 2017). The significant time spent in an escalating series of meetings was useful initially, but after relevant officials absorbed the policy and its application, some streamlining might have been in order. Still, the diversity of perspective that accompanied these decisions added value. In addition to bringing broader concerns into the discussion, the presence of possibly skeptical officials in the discussion requires proponents of an operational plan to develop their own arguments more carefully, thereby identifying and addressing flaws.[2] The Trump administration's short-circuiting of interagency review removed those benefits.

Finally, the lack of strong external oversight from the judiciary or even Congress undermines accountability for these operations. There are practical

constraints on how widely or frequently the executive branch is able to share information about these strikes. Information about individual operations can be sensitive, particularly in advance of action. In addition, the facts underlying these operations change frequently, and decisions often must be made quickly. Inserting external oversight in the middle of that decision-making can decrease effectiveness. Nonetheless, external oversight, even if after the fact, is an important way to improve the quality of analysis and decision-making. It can do this directly by identifying and correcting errors and missteps and holding operators accountable for mistakes. The anticipation of external oversight also can enhance quality and effectiveness indirectly by forcing decision-makers to more carefully develop, articulate, and defend their policies and operations.[3]

2.3 Possible Reforms: Target-Designation Phase

Despite these concerns, finding workable measures to increase accountability for operational planning for lethal targeting is challenging.

Some have proposed creating a mechanism for prior judicial review of lethal targeting decisions – perhaps a special court, similar to the Foreign Intelligence Surveillance Court (FISA Court), which operates *ex parte* to consider applications for electronic and physical surveillance for foreign intelligence purposes (Guiora & Brand 2015; *New York Times* 2013; Shane 2013). These proposals have never gotten far. The practical hurdles for this type of review are significant. The intelligence on which targeting decisions are based is fluid and the decision-making is dynamic. If court review comes before a strike opportunity, the facts on which the court was asked to base its decision could be outdated by the time a strike becomes feasible. If the court is asked to rule when the target is in sight, its ability to consider the circumstances fully will be constrained. Requiring the president to wait for court approval before ordering a military operation runs the risk of infringing on the president's constitutional role as commander-in-chief (Vladeck 2013; Shane 2013).

From the perspective of the courts, this role would likely be uncomfortable and perhaps unconstitutional. US courts typically engage in *ex post* review, not *ex ante*, and their proceedings are adversarial, not *ex parte*. There are exceptions to this; most relevant here is the courts' power to issue search warrants. But when courts issue search warrants, there is an anticipated future adversarial proceeding.[4] This prospect of an adversarial proceeding would not exist for an *ex parte* order that approves a target for a lethal strike. This raises significant rule of law and even constitutional issues (Vladeck 2013, 2014, fn. 44).[5] In fact, judges have not appeared eager to take on this responsibility. As one retired judge said, it is "not the business of judges…to decide without an adversary party to sign a death warrant for somebody" (Shane 2013).

Increased congressional review of pre-strike decision-making is a more promising option and could impose beneficial discipline on executive branch decision-making. The challenge here is that Congress does not control the resources or staffing of the executive branch. It is difficult for members of Congress and their staffs to develop the deep expertise needed to provide effective, systematic oversight of these classified operational proposals. Although members of Congress and their staffs can nonetheless add real value to the process by asking tough questions of the executive branch, the political incentives tend to run against significant involvement for many legislators. Members of Congress may see little upside to committing precious time to an issue that does not directly affect the lives of their constituents and only generates publicity when something goes wrong.

Despite its limitations, improved executive branch process and standards may be the most practical step to increase accountability for lethal targeting operations. Improved executive branch process could streamline the PPG decision-making, but maintain critical interagency perspectives for all operational plans. The addition of a red-team or devil's advocate process, at least for some more difficult or controversial decisions, could assure more informed final decisions. Although not every operational plan requires multiple levels of meetings, approval at a level above the entity carrying out the operation would ensure more rigorous consideration of each proposal.

Additional small steps could contribute to better decisions, increased accountability, and a more credible process. Requiring clear and consistent record keeping about targeting decisions would contribute to a more disciplined process and facilitate accountability after the event. Increasing clarity about standards and transparency about how they are applied would enhance the credibility of targeting decisions with the US public and internationally. Imposing these types of process changes by legislation, rather than by executive branch policy, would make them more durable.

3. OPERATIONAL PHASE

3.1 Operational Details

This section describes how targeted killing operations are conducted, identifies the risks that they pose, and analyzes internal and external mechanisms designed to ensure accountability for them.[6]

The current aircraft used by the US is the Reaper, which has a 60-foot wingspan, a maximum speed of 300 miles per hour, and a flying altitude of 50 000 feet. It can stay airborne for 16–18 hours and may carry two 100-pound supersonic hellfire missiles and two 500-pound laser-guided bombs. Its video functions are based on technology known as Gorgon Stare, which provides

wide-area surveillance capable of capturing motion imagery of an entire city. However, a slight time lag before imagery is transmitted means that the technology is not used to conduct the strike itself.

The aircraft is operated from a ground control station (GCS) whose crew includes:

1. a pilot, who controls the flight of the aircraft and fires weapons;
2. a sensor operator, who gathers information from sensors on the aircraft and video surveillance reports to guide the flight, track ground objects, mark targets with lasers, and guide the weapon once released to the target;
3. a mission intelligence coordinator, who monitors intelligence sources outside the GCS and serves as liaison between the cockpit and the outside; and
4. a safety observer, who enters a GCS when the possibility of firing a weapon arises.

A launch and recovery element (LRE) team near the location of the aircraft launches it and turns control over to the GCS. The GCS controls the aircraft during a strike, and then transfers control to the LRE to return it to base. Each mission is supported by up to 200 people. There are people outside the GCS involved in flying the aircraft, multiple sources of airborne and ground intelligence for each mission, and scores of intelligence analysts.

Target identification relies on intelligence sources to direct the aircraft to the target. There must be positive identification (PID) of the target by at least two sources, and an unbroken view of the target at all times, or the PID process must start over. An authorization to conduct a strike specifies coordinates of the target, the location of any friendly forces, and the intent of the ground commander.

Following authorization, there is a pre-strike collateral damage estimate (CDE) of anticipated civilian casualties. A pilot is then "cleared hot" to fire from an optimal launch altitude generally of 10 000 feet. The weapon leaves the aircraft and seeks the lasered target, guided to it by the sensor operator, who keeps the guidance system crosshairs on the target to ensure that the missile hits it. If civilians unexpectedly appear, the sensor operator has about 30 seconds to shift the laser pointer to a pre-designated area that will avoid casualties.

After the strike, there is a battle damage assessment (BDA), which assesses accomplishment of the mission; damage to military, civilian, and dual-use objects; and hostile and innocent casualties resulting from the strike. There is a detailed video and computer audit trail of every mission, which is used for a debriefing of what went well and what did not. Lessons are incorporated into

training, and a Remedial Action Report identifies errors, which can result in persons being removed from missions while they undergo retraining.

3.2 Accountability Concerns

There are risks with any use of force, but some are distinctive to drone operations that may result in striking the wrong target or causing civilian casualties. Several steps in the targeting process are intended to minimize these risks and to provide accountability for these operations.[7]

First, a strike relies on a complex system of processing, exploiting, and disseminating information among multiple participants in an extended chain. The GCS receives and synthesizes inputs from numerous imagery analysts to gain the best possible situational awareness. Imagery analysts, however, do not have any direct contact with anyone on the ground; only the GCS crew does. This can create the risk of incomplete situational awareness in operations in which ground forces play a role.

Second, a video feed cannot identify individuals in a crowd, so operators must rely on other sources of intelligence for PID. Continuous PID can be challenging as individuals go into buildings or otherwise disappear from the video feed. When they do so, the process of confirming PID based on two independent sources must be reinitiated. This can serve as an internal accountability mechanism, but must rely in many cases on sources of intelligence other than the video feed.

The requirement of a CDE prior to a strike is another accountability mechanism. Estimates are based, however, not on historical experience with similar operations in the area, but on models that include data on weapons characteristics and environmental conditions (Sewall 2017). This can limit the extent that estimates are sensitive to the social environment, although pattern-of-life analysis can help compensate for this limitation.

Identifying civilians also can be challenging. Intelligence might not be available on persons in the target area who are not the target. In that case, analysts must infer from ambiguous appearance and behavior who is a civilian and who is a militant. In addition, video feeds cannot see into buildings, which means that they may not show civilians in them.

The targeting process begins with a wide video focus and then zooms in right before a strike. This can create a "soda straw" effect that limits awareness of last-minute entry of civilians into the target area. Gorgon Stare provides expansive video awareness, but the slight delay in its feed means that it cannot be used for real-time overwatch and strike prosecution.

The post-strike BDA also can serve further accountability by verifying the identity of casualties. There are challenges, however, in relying on a BDA for information on civilian casualties. First, the traditional purpose of a BDA has

not been to identify such casualties, but to determine if targets were successfully destroyed. Expanding the BDA to serve other purposes requires not only a change in procedure but a change in assessment perspective.

In addition, a drone is used in some cases because the target is in a remote area where the US has no ground forces and few, if any, sources of intelligence. In these cases, a BDA will be based solely on assessment by aerial assets. There are limits, however, to the information that these assets can provide. Video of building rubble, for instance, will not necessarily indicate if there are bodies underneath it. This means that a BDA may not accurately determine the number of casualties and whether they are militants or civilians. The latter information typically will require local sources of information unavailable from the air.

This raises a related point. Some observers claim that there are significant deficiencies in post-strike investigations even when they go beyond reliance on aerial assets (Center for Civilians in Conflict [CIVIC] 2020; Crawford 2013, p. 89; Lewis & Holewinski 2013). One criticism is that investigations do not seek local sources of information, such as humanitarian organizations working in the area, residents, and local social media. One partially declassified study conducted by the National Defense University at the direction of the Chair of the US Joint Chiefs of Staff found that 58% of information on civilian casualties from 2015 to 2017 came from outside sources (US Joint Chiefs of Staff 2018, p. 17). There is a tendency, however, to give credence only to those reports that can be confirmed by internal military assets, such as video.

A study of military civilian casualty investigations by CIVIC also found that "[c]ivilians, civil society, and others often face barriers when trying to make complaints of civilian harm to the military" (CIVIC 2020, p. 3). This and other problems with investigations led CIVIC to conclude that "over the last eighteen years, examples of good practice in investigating civilian harm have been overshadowed by the inconsistency—and, too often, inadequacy—of the overall record of military investigations" (CIVIC 2020, p. 1).

These deficiencies undermine the critical foundation of any effort to assure accountability: that the US fully understands the consequences of those operations. It is important not only to conduct effective investigations, but to assemble data from them that permits root cause analysis of the causes of casualties. The International Security Assistance Force in Afghanistan established a civilian casualty tracking cell in 2008 to standardize reporting procedures, but generally was unable to conduct root cause analysis (CIVIC 2014). Sara Sewall and Larry Lewis undertook such analysis in Afghanistan in 2010 with support from the military, but this has not occurred in other theaters (Sewall & Lewis 2010; US Joint Forces Command 2011).

The debriefing after a strike also can further accountability, but its effectiveness is limited by the potential gaps in information about the consequences of

a strike that we have described. In addition, there is no effort to systematically incorporate these into a larger database that would enable analysts to identify root causes.

In addition to the potential internal sources of accountability we have discussed, external ones are available. We review these in the next section.

3.3 External Accountability

Strikes by the military must be reported to the congressional defense committees by the Secretary of Defense within 48 hours (10 USC §130f). The Armed Services Committees have responsibility for a wide range of military issues, however, and not all members may have the expertise or security clearance to fully assess these activities.

Each strike conducted by any other agency is reported to the congressional intelligence committees, which also must be notified of any strike conducted as a covert action by any agency (10 USC §3093). These committees generally are well regarded, but it is unclear whether they receive sufficient details to assess whether PID and civilian casualty estimates are accurate for each strike.

Disclosure of civilian casualties also can be an important source of accountability. In 2016, President Obama issued an executive order requiring the Director of National Intelligence (DNI) to provide an annual public report on the number of strikes by the US outside areas of active hostilities, including information on civilian casualties. The order required that this report identify the sources of information and methodology used to determine this number. The order also directed the DNI to address reasons for "discrepancies between post-strike assessments from the U.S. Government and credible reporting from nongovernmental organizations regarding non-combatant deaths" (White House 2016a). Congress has reiterated these requirements in its appropriations process.

The reference in the executive order to nongovernmental organizations reflects investigations by several such organizations over the years that have documented civilian casualties, notwithstanding US claims to the contrary (Open Society 2015; Ross 2014; Human Rights Watch 2013; CIVIC 2010). These rely on sources of information that the government has not consulted or obtained through local interviews or social media. Such sources can provide first-hand accounts, particularly regarding the plausibility of characterizing individuals as militants or civilians. Not all sources are necessarily accurate, since memories may not be fully accurate, and some sources may have reasons for describing events in a certain way. These investigations nonetheless have provided valuable information and have served as an impetus for the US to be more forthcoming about civilian casualties resulting from strikes.

In the same vein, Congress has required the Department of Defense to establish an office to coordinate efforts to minimize civilian casualties, including establishing uniform processes for investigating such casualties and developing best practices for reducing casualties. This office is engaged in preparation of a civilian casualty policy that will be applicable across the Department (Undersecretary of Defense 2020).

3.4 Possible Reforms: Operational Phase

As we have described, various internal and external mechanisms are available that can contribute to accountability for targeted strikes. What additional measures might enhance such accountability?

Most important is improving the investigative process so that the US has an accurate understanding of the impact of strikes, and that it uses the results to analyze the reasons for civilian casualties. Expanding outreach to nongovernmental and local sources of information would be an important step in this direction, especially when the US has no ground assets in a strike location. This would strengthen the post-strike BDA, provide information that would improve the After Action Review, enhance training for targeted operations, and could generate data that would allow the CDE to be based on historical experience to some extent.

Greater disclosure of information about individual strikes to the public and the congressional committees, consistent with protecting intelligence sources and methods, also could improve accountability. This might include a general description of the basis for targeting an individual and why capture was infeasible. Disclosure also could include information on discrepancies between the CDE and the actual number of civilian casualties, with a good-faith attempt to explain the reasons for the difference. Such detail is especially important for covert strikes, which the government generally does not disclose to the public. It also is important for congressional staff to receive sufficient training to be able to effectively scrutinize this information.

Finally, there has been debate about whether targeting should be conducted only by the Defense Department. Some observers have suggested this on the grounds that the military has greater experience operating under rules governing the use of force. To the extent that both the military and other agencies continue to conduct strikes, it may be worth considering eliminating bifurcated congressional review by establishing a joint congressional subcommittee with oversight authority for all targeted strikes. This could enable oversight by a body that develops expertise on such operations and therefore is able to conduct a searching inquiry into them.

4. CONCLUSION

Targeted killing operations pose especially difficult challenges for governmental accountability in liberal democracies. Disclosure of criteria for targeting, robust and expansive internal deliberation, thorough investigation of the impact of strikes that draws on several sources of information, and reports on individual strikes that are as detailed as possible hold some promise in meeting these challenges. The tension between the perceived need for secrecy and the momentous nature of these operations nonetheless is likely to result in ongoing debates about accountability.

NOTES

1. See Gil (2020, p. 729), who discusses the reluctance of US Courts to become involved in targeted killing decisions: "As for the courts, several lawsuits were filed in connection with targeting decisions but all were dismissed on various grounds. The opinions stressed, inter alia, that courts have no part to play in this area, both because 'judicially discoverable and manageable standards' for adjudication are lacking and because military decisions should be 'in the hands of those who are best positioned and most politically accountable for making them.'" See Alston's (2011, pp. 393–402) analysis of how courts have invoked the political question, standing, and state-secret doctrines to avoid ruling on the substance of claims involving targeted killings; *Al-Aulaqi v. Obama*, 727 F. Supp. 2d 1 (D.D.C. 2010), which dismisses, on political question and standing grounds, claims regarding the alleged designation of a US citizen for targeting; and *Al-Aulaqi v. Panetta*, 35 F. Supp. 3d 56 (D.D.C. 2014), which dismisses a damages suit arising from the killing of two US citizens in a targeted strike.
2. See Chambers (2004, p. 391), which explains that the requirement to justify oneself to others creates "the necessity to articulate one's position carefully, to defend it against unexpected counter arguments, to take opposing points of view into consideration, to reveal the steps of reasoning one has used, and to state openly the principles to which one appeals."
3. See Chambers (2004), as well as DeRosa and Regan (2018, p. 32), who discuss the risks that limiting external oversight can pose: "If there is minimal likelihood that other branches or the public will demand a full well-reasoned explanation for a decision, the quality of the deliberation that precedes it may suffer. This in turn can impair its perceived legitimacy."
4. FISA Court *ex parte* warrants are justified on the grounds that they, like search warrants, contemplate the availability of a future adversarial proceeding (Vladeck 2013).
5. As Vladeck (2013) explains: "The Supreme Court has long emphasized, as it explained in *Flast v. Cohen*, that one of the central purposes of Article III's 'case-or-controversy requirement' [in the US Constitution] is to ensure that 'the dispute sought to be adjudicated will be presented in an adversary context and in a form historically viewed as capable of judicial resolution.' That is to say, 'adversity' is one of the cornerstones of an Article III case or controversy, and it would be noticeably lacking in a drone court."

6. The operational details of missions described in this section are drawn from Lee (2019), Martin (2010), and Woods (2015).
7. Unless otherwise indicated, the details in this section are drawn from work by Larry Lewis at the Center for Naval Analyses (CNA), a federally funded research and development center that provides research and analysis services to military and government agencies on issues relating to national defense, and an interview with Lewis on February 18, 2020. The CNA has the advantage of access to information that many other research entities do not, and its work is rigorous and very highly regarded by the wide range of actors who work and are involved with national security issues (although its work, of course, does not disclose classified information). Publications that form the basis for the discussion in this section on which Lewis has worked are Lewis and Varichek (2016), Lewis (2014), Lewis and Holewinski (2013), Center for Army Lessons Learned (2012a, 2012b), and Sewall and Lewis (2010).

REFERENCES

Alston, Philip (2011), 'The CIA and Targeted Killings beyond Borders', *Harvard National Security Journal*, 2(2), 283–446.

Atherton, Kelly (2020, May 22), 'Trump Inherited the Drone War but Ditched Accountability', *Foreign Policy Online*, accessed November 19, 2020, at https://foreignpolicy.com/2020/05/22/obama-drones-trump-killings-count/.

Center for Army Lessons Learned (2012a), *Improving Lethal Action: Learning and Adapting in U.S. Counterterrorism Operations.* Arlington, VA: CNA & Center for Army Lessons Learned.

Center for Army Lessons Learned (2012b), *Afghanistan Civilian Casualty Prevention*. Fort Leavenworth, KN: Center for Army Lessons Learned.

Center for Civilians in Conflict (CIVIC) (2010), *Civilians in Armed Conflict: Civilian Harm and Conflict in Northwest Pakistan*, Research Report. Washington, D.C.: CIVIC.

Chambers, Simone (2004), 'Behind Closed Doors: Publicity, Secrecy, and the Quality of Deliberation', *The Journal of Political Philosophy*, 12(4), 389–410.

CIVIC (2014), *Civilian Harm Tracking: Analysis of ISAF Efforts in Afghanistan*, Research Report. Washington, D.C.: CIVIC.

CIVIC (2020), *In Search of Answers: U.S. Military Investigations and Civilian Harm*, Research Report. Washington, D.C.: CIVIC.

Crawford, Neta (2013), *Accountability for Killing: Moral Responsibility for Collateral Damage in America's Post-9/11 Wars*. Oxford: Oxford University Press.

DeRosa, Mary & Regan, Mitt (2018), 'Deliberative Constitutionalism in the National Security Setting', in Ron Levy, Hoi Kong, Graeme Orr, & Jeff King (eds.), *The Cambridge Handbook of Deliberative Constitutionalism*. New York: Cambridge University Press, pp. 28–43.

Gil, Elad D. (2020), 'Institutional Choice and Targeted Killing: A Comparative Perspective', *Tulane Law Review*, 94(4), 711, 729.

Guiora, Amos & Brand, Jeffrey (2015), 'Establishment of a Drone Court: A Necessary Restraint on Executive Power', in Steven J. Barela (ed.), *The Legitimacy of Drones*. Farnham: Ashgate, pp. 323–58.

Hartig, Luke & Tankel, Stephen (2019, August 15), 'Part II: The Muddy Middle: Challenges of Applying Use of Force Policy Guidance in Practice', *Just Security*,

accessed March 26, 2021, at https://www.justsecurity.org/65819/part-ii-the-muddy-middle-challenges-of-applying-use-of-force-policy-guidance-in-practice/.
Human Rights Watch (2013), *'Between A Drone And Al Qaeda': The Civilian Cost Of US Targeted Killings in Yemen*. New York: Human Rights Watch.
Lee, Peter (2019), *Reaper Force: The Inside Story of Britain's Drone Wars*. London: John Blake.
Lewis, Larry (2014), *Improving Lethal Action: Learning and Adapting in U.S. Counterterrorism Operations.* Arlington, VA: CNA & Center for Army Lessons Learned.
Lewis, Larry & Holewinski, Sarah (2013), 'Changing of the Guard: Civilian Protection for an Evolving Military', *Prism*, 4(2), 57–66.
Lewis, Larry & Varichek, Diane (2016), *Rethinking the Drone War.* Quantico, VA: CNA & Marine Corps University Press.
Martin, Matt (2010), *Predator: The Remote-Control Air War over Iraq and Afghanistan: A Pilot's Story.* London: Zenith Press.
McNeal, Gregory (2014), 'Targeted Killing and Accountability', *Georgetown Law Journal*, 102, 681–794.
New York Times (2013, February 13), 'Editorial: A Court for Targeted Killings', accessed March 26, 2021, at https://www.nytimes.com/2013/02/14/opinion/a-special-court-is-needed-to-review-targeted-killings.html.
Open Society (2015), *Death By Drone: Civilian Harm Caused by U.S. Targeted Killings in Yemen*. New York: Open Society Foundations.
Ross, Alice (2014), 'Leaked Official Document Records 330 Drone Strikes in Pakistan', *The Bureau of Investigative Journalism,* accessed March 26, 2021, at https://www.thebureauinvestigates.com/stories/2014-01-29/leaked-official-document-records-330-drone-strikes-in-pakistan.
Savage, Charlie & Schmitt, Eric (2017, September 21), 'Trump Poised to Drop Some Limits on Drone Strikes and Commando Raids', *New York Times*, accessed March 26, 2021, at https://www.nytimes.com/2017/09/21/us/politics/trump-drone-strikes-commando-raids-rules.html.
Sewall, Sarah B. (2017), *Chasing Success: Air Force Efforts to Reduce Civilian Harm*. Maxwell Air Force Base, AL: Air University Press.
Sewall, Sarah B. & Lewis, Larry (2010), *Joint Civilian Casualty Study Executive Summary*. Washington, D.C.: US Department of Defense.
Shane, Scott (2013, February 8), 'Debating a Court to Vet Drone Strikes', *New York Times*, accessed March 26, 2021, at https://www.nytimes.com/2013/02/09/world/a-court-to-vet-kill-lists.html?_r=0.
Undersecretary of Defense (2020, January 31), *Development of a DoD Instruction on Minimizing and Responding to Civilian Harm in Military Operations*. Washington, D.C.: U.S. Department of Defense.
US Joint Chiefs of Staff (2018, April 17), *Executive Summary Civilian Casualty (CIVCAS) Review*. Washington, D.C.: U.S. Department of Defense.
US Joint Forces Command (2011), *Adaptive Learning for Afghanistan*. Suffolk, VA: Joint Center for Operational Analysis.
Vladeck, Stephen (2013, February 10), 'Why a "Drone Court" Won't Work – But (Nominal) Damages Might…', *Lawfare Blog*, accessed March 26, 2021, at https://www.lawfareblog.com/why-drone-court-wont-work-nominal-damages-might.
Vladeck, Stephen (2014), 'Response: "Targeted Killing and Judicial Review"', *George Washington Law Review*, 82, 11.

White House (2013, May 22), 'Procedures for Approving Direct Action against Terrorist Targets Located outside the United States and Areas of Active Hostilities', accessed March 26, 2021, at https://www.justice.gov/oip/foia-library/procedures_for_approving_direct_action_against_terrorist_targets/download.

White House (2016a), 'Executive Order 13732—United States Policy on Pre- and Post-Strike Measures to Address Civilian Casualties in U.S. Operations Involving the Use of Force', accessed March 26, 2021, at https://obamawhitehouse.archives.gov/the-press-office/2016/07/01/executive-order-united-states-policy-pre-and-post-strike-measures.

White House (2016b), *Report on the Legal and Policy Frameworks Guiding the United States' Use of Military Force and Related National Security Operations*. Washington, D.C.: The White House.

White House (2020), *Report on the Legal and Policy Frameworks Guiding the United States' Use of Military Force and Related National Security Operations*, accessed March 26, 2021, at https://www.state.gov/wp-content/uploads/2019/10/Report-to-Congress-on-legal-and-policy-frameworks-guiding-use-of-military-force-.pdf.

Woods, Chris (2015), *Sudden Justice: America's Secret Drone Wars*. Oxford: Oxford University Press.

6. Interrogation ethics in counter-terror operations

Michael Skerker

1. INTRODUCTION

This discussion of interrogation ethics in the context of counter-terror operations will begin by sketching out the normative framework for identifying the norms that interrogators must observe (Sections 2 and 3).[1] We will then address the foundation of interrogators' duties (Section 4) and the moral rights of people suspected of terrorism offences (Section 5) before detailing the most norm-compliant interrogation techniques (Sections 6 and 7). The chapter will focus on moral rights and permissions rather than legal rights and liberties. Restrictions on space preclude addressing the right modes of interdiction, detention, and trial of terrorism suspects. This chapter will not address interrogational torture since torture is not a reliable method of gaining Intelligence; the chapter will instead focus on non-coercive interrogation methods (see Skerker 2010).

2. PROFESSIONAL NORMS

Professions are permitted to exist as semi-autonomous, self-regulating sub-cultures in society because they are dedicated to the promotion or protection of basic goods that all people need for a flourishing life, including health, education, justice, and safety (Camenisch 1983, pp. 54–5). Part of what identifies professions as professions is their internal promulgation of professional norms—norms guiding professionals to efficiently deliver the signal goods of the profession with a minimum of social disruption. These norms develop over time based on expert opinion from professionals about what meets the professions' ends, and lay opinions from the professionals' clients.

In adversarial professions, where professionals protect clients' interests against the interests of some non-clients, professional norms take into account the rights and interests of the three stakeholder groups relevant to professional activity: the professionals themselves, their clients, and non-clients.

Elsewhere, I have argued that norms are identified by a winnowing process called the security standard in which the most rights-respecting out of the most practically efficacious norms and tactics are selected (see Skerker 2020, Ch. 5).

Norms can be thought of as rules for selecting tactics. While norms can be context independent, feasible tactics will depend on local context. When applying the security standard, one first considers which locally feasible norms and tactics are the most reliable, proportionate, efficient, and effective at accomplishing the profession's signal goals. One then asks which among the contenders best respects the rights of those in the three stakeholder groups. The rights element is itself a balancing of the rights of members of the three groups, animated by a principle of reciprocity. We will consider below how different non-coercive interrogation methods fare by this standard. Here, it is worth showing how poorly interrogatory torture fares by the standard. Torture is an unreliable method for garnering Intelligence because the interrogator cannot trust anything the torture victim says—the interrogator, with their torments, has provided the victim ample reasons to lie. For the same reason, and because of the amount of false information it invariably produces, torture is also not an effective interrogation method. The amount of harm done to the victim is disproportionate to the amount of reliable and truthful information typically produced. The amount of false or irrelevant information produced alongside the amount of true information produced—but often unacknowledged by the interrogators, who do not know what interrogators do not know—also indicates the inefficiency of torture. Finally, torture grossly violates the victim's rights, as many consider the right not to be tortured to be absolute. Due to its relative inefficacy, inefficiency, and unreliability, torture performed by state agents also fails to protect the rights of their citizens to security. Torture also tends to damage interrogators' psyches to such an extent (Arrigo 2004, pp. 543–72; Lagouranis 2008; Phillips 2012; Edmonds 2015) that it seems that their rights are violated if forced by the state to employ such measures (Skerker 2019).

3. NORM-COMPLIANT INTERROGATION

In the case of interrogation, be it in law enforcement, military, or Intelligence contexts, the practical elements of the security standard serve to protect the rights and interests of interrogators' clients—the inhabitants of their city or state—by garnering true confessions and/or actionable Intelligence. Innocent suspects are also benefited by reliable and proportionate techniques that will hopefully reveal their innocence with a minimum of stress and anxiety. In a material sense, guilty suspects are not benefited by effective interrogations—they might well prefer interrogations by incompetent interrogators—but their potential preferences to remain unpunished for their crimes are not morally

relevant preferences that need to be taken into account by professional norms. Even guilty suspects are benefited by practically effective techniques insofar as the approaches do not cause more distress that can be proportionally justified, they will not lead guilty suspects to confess to worse crimes than they actually committed, and suspects will not be subjected to ineffective and potentially onerous techniques. Interrogators are benefited by effective techniques in both law enforcement and national security contexts insofar as they are members of the public who benefit from crime suppression, or, in military contexts, insofar as they are members of a military who benefit from obtaining tactical information about the enemy. The use of practically effective approaches also mitigates legal risk and moral injury to those involved in the interrogation.

Regarding the rights-respecting element of the security standard, interrogators' clients' rights to safety are largely met through the use of effective norms and tactics. Innocent and guilty suspects interrogated in norm-compliant ways have their rights respected to a degree consistent with effective interrogation. While the practical and rights elements of the security standard are in tension—leading one to privilege different norms or tactics than one might in the absence of the opposing consideration—they are not opposed. Effective interrogation norms and techniques infringe on some rights, such as the rights to privacy and to honest dealing, but come as part of a general law enforcement enterprise designed to protect people's rights. The rights of interrogators are respected by norm-compliant interrogation techniques insofar as they are potentially the target of interrogators if they themselves are ever suspected of a crime, or, in military contexts, if they are captured by the enemy. Their rights are also respected if they are not required to do things that jeopardize their psychological health or character.

4. THE FOUNDATION OF PROFESSIONAL DUTIES

The security standard identifies professional norms but does not justify their use. Professionals have moral duties to adhere to their security-standard-compliant norms because of a collective moral responsibility to respect others' rights (Skerker 2020, Ch. 5). Collective moral responsibilities are extensions of the duties all have to respect others' rights and to help protect the goods necessary for rights enjoyment. Collective moral responsibility is attached to individuals when they live in groups because certain rights, such as the right to safety, can only optimally be met over time through coordinated collective action (Miller 2010, Ch. 4). Most people meet their collective moral responsibilities by supporting institutions. They have a moral duty to support just institutions that expertly, reliably, and impartially protect the rights of people over time. Professionals in these institutions meet their relevant collective moral responsibilities by adhering to their properly constituted professional norms that

are identified by the security standard. Thus, their moral duty to support just institutions transforms into a moral duty to adhere to their properly constituted professional norms (Miller 2010, Ch. 2).

5. THE RIGHTS OF SUSPECTS IN COUNTER-TERROR INTERROGATIONS

While interrogators have duties to extract information from their targets, they must also omit some potentially efficacious actions in deference to suspects' rights. This section will identify the relevant rights of interrogation targets in counter-terror contexts. The relevant persons interrogated might be conventional combatants, privileged irregulars, or unprivileged irregulars. Conventional combatants are members of the armed forces. According to conventional Just War Theory as well as the law of war, combatants' political leaders are accountable for the decision to go to war; combatants are legally permitted to obey deployment orders and fight their opponents irrespective of the justice of the cause (Walzer 2015, pp. 34–41).

Non-state actors in organized armed groups can enjoy the same belligerent privileges provided certain tangible connections with a nascent or extant political entity and adherence to the laws of war. Even though they may lack some of the features of conventional combatants, these irregular militants are more like soldiers than criminals (Walzer 2015, pp. 179–86).

Militants who lack these connections and/or fail to adhere to the laws of war are 'unprivileged irregulars'. They are effectively criminals. They may well be motivated by political ideals, but they have taken it upon themselves to fight for these ideals without the justifications that come from concrete relationships with specific political communities. In fact, it is more likely that these militants force their esoteric ideas on populations they coerced into submission (see Skerker 2010, Chs. 2 and 6; Skerker 2011).

I use the term 'unprivileged irregular' militant to refer to those persons whom many call 'terrorists' (see Skerker 2010, Ch. 6). I do not think the term 'terrorist' is very useful since the term identifies a person by the tactics that terrorists use rather than by their political status. Individuals and members of states and a variety of sub-state groups can engage in terrorism and intentionally target non-combatants and their infrastructure for political effect.

Privileged combatants (both conventional and irregular) have a right to their tactical secrets. It is permissible for them to keep this information secret because it is pursuant to legitimate activities. Unprivileged irregulars do not have a right to tactical secrets since these secrets are pursuant to rights-violating activities. Still, unlike innocent civilians, none of these groups have cause to complain if adversary state agents attempt to elicit information they might not want to share. The reason unprivileged irregulars lack this priv-

ilege is clear. For their part, privileged combatants are permitted to engage in adversarial behaviour in defence of their state's or their community's security, including, but not limited to, violence. Various methods of Intelligence gathering, including interrogation, are permitted as a way of countering the enemy's war efforts. To draw an analogy, poker players have a right to their money, but their opponents do not wrong them by bluffing and engaging in other kinds of strategic behaviour to get their money from them.

Still, there are limits that interrogators must observe with both privileged and unprivileged combatants. Since privileged combatants have a right to their tactical secrets, they cannot be threatened or otherwise coerced into giving them up. These actions would only be, in principle, permitted if prisoners of war (POWs) were wronging their captors or their citizenry by holding them. Still, captors can trick, control, or offer inducements to POWs for their secrets. In the case of threats, blackmail, or coercion, the POW's autonomy is not respected. In the case of inducements, the POW's autonomy is respected because it is up to the POW whether or not to accept the enticement. In cases of interrogators' strategic behaviour, such as tricking or emotionally manipulating the POW into revelations, the POW is engaged as a strategic opponent; the POW is not treated with consummate respect for their autonomy since the interrogator tries to shape the prisoner's understanding of reality, but the latter is free to refuse participation or to attempt to deceive the interrogator in turn (see Skerker 2010, Chs. 6 and 7). To return to the poker analogy, a player can attempt to bluff, but cannot simply grab their opponent's poker chips or threaten their opponent with bodily harm in order to get them.

In principle, it *would* be permissible to threaten, blackmail, or use proportionate levels of force to get unprivileged irregulars to reveal tactical secrets since they have temporarily forfeited a right to them, and their concealing of these secrets threatens innocent people. However, in many cases, the unprivileged irregular is not positively identified as an unprivileged irregular by being caught while perpetrating violent actions, wearing some sort of uniform (as members of the Islamic State [ISIS] sometimes do, for example), or by self-identifying. Rather, in many tactical situations, the suspected member of al-Qaeda, ISIS, al-Shabaab, and so on is a person in civilian clothing, vigorously protesting their innocence. As a *suspected* unprivileged irregular, the individual is effectively a criminal suspect, someone who may well be an innocent person with an intact right to privacy and silence. While the reader may be thinking of criminals as domestic criminals (for example, bank robbers or drug dealers), an unprivileged irregular's violence is criminal on account of not being morally and legally privileged in the sense that a conventional combatant's wartime activities on behalf of a state are privileged. The unprivileged irregular killing people or destroying property is violating both the domestic law of the state and international law.

A moral right to (mental) privacy is important to express and defend one's autonomy. This right gives one power to protect one's thought and deliberation. Since knowledge of others' thoughts, values, memories, opinions, and self-knowledge (for example, of past actions) can give one power over the target, a right to privacy includes the power to control the release of this kind of personal information (see Skerker 2010, Ch. 3). A right to silence is an expression of the right to privacy when someone asks personal questions. This right means one does not wrong strangers if one refuses to answer personal questions; they are wrong to try to force one to speak.

Like many rights, a right to privacy can be abused, giving others a right to justifiably infringe on that right's exercise. For example, one would be morally permitted to seize a golf club from someone who refused to stop swinging it on a crowded train car. Even if the attacker was not targeting a particular person, their reckless use of their property at least temporarily forfeits their claim to it. The proper extent of a person's rights exercise can be modelled by querying the scope of exercise consistent with universal exercise. Rights violations or threats to rights usually cannot be universalized (see Skerker 2010, Ch.1). For example, the golfer could not rationally consent to everyone else swinging their golf clubs on the train at the same time, as such activity would prevent them from taking a swing. While one's mere thinking is consistent with all others' thinking their own thoughts, and keeping one's personal secrets is consistent with everyone else keeping their secrets, universal plotting or universal concealment of criminal information is not universalizable if such mental activity is considered part of the physical rights violations that plotting precedes or concealment follows. Another way of putting the point is that we do not defer to others' mental privacy, giving them the space to formulate their own plans and decide what information to share with the world, for them to plot our demise or protect murderers from punishment.

An unprivileged irregular who refuses to divulge tactical information would not be wronged by proportionate treatment designed to force them to provide the information, such as blackmail or threats of permissible consequences like imprisonment or deportation. Yet, again, a *suspected* unprivileged irregular might be an unprivileged irregular or might be an innocent person. Given this ambiguity, it would be wrong for interrogators to assume a suspect has forfeited their right to privacy and would not be wronged by threats or blackmail. Since a suspected unprivileged irregular might be innocent, an interrogator must respect as many rights that innocent people normally enjoy, consistent with what all would consent to as a means of identifying unprivileged irregulars (more on this below). Relevant general rights include the right to honest dealing, the right to privacy (and silence), and the right to be treated with respect (for example, not to be humiliated, demeaned, cursed, shouted at, and so on). If the detaining power plans to prosecute the suspected unprivileged

irregular, interrogators must also respect rights particular to the criminal justice arena, namely, due-process rights. These rights provide a suspect a chance to assert their innocence and maintain a degree of autonomy, particularly with respect to the decision to reveal their secrets or not. Due-process rights do not simply benefit the suspect. Reminding them about their privilege against compelled self-incrimination and affording them a right to challenge witnesses, see inculpatory evidence, hire an attorney, and remain silent in interrogation and trial without judicial prejudice also increases the reliability of any confessions and subsequent verdicts (see Skerker 2010, Chs. 3 and 4). Investigators have to amass a compelling degree of evidence if they cannot simply beat a confession out of the suspect, and prosecutors have to also mount a robust case if the defence has chances to challenge the state's case.

Just the same, interrogators cannot treat suspects simply as they would an apparently innocent person they saw on the street. In that case, interrogators would not be interrogating them at all. An interrogator must treat suspects somewhat strategically to ascertain their guilt or innocence. Respecting the suspect's right to remain silent is key to integrating the opposing pressures on the interrogator (to treat the suspect as potentially innocent and potentially guilty). Innocent people have a right to their personal information and do no wrong by refusing to answer strangers' probing questions. They also have a right to physical liberty and to honest dealing. Interrogators cannot do their job without physically detaining suspects. Effective gambits (see below) may include deceptive and emotionally manipulative elements. Again, suspects may be guilty and, so, may lack a right to refuse cooperation with authorities. Thus, a suspected unprivileged irregular may be treated worse than a suspect deserves, if the suspect is innocent, and better than the suspect deserves, if the suspect really is an unprivileged irregular. Respecting the right to silence is a way of acknowledging the suspect's ambiguous status and operationalizing the interrogator's contrary imperatives (to respect the rights of the innocent and to pursue unprivileged irregulars). While the suspect is not free to leave and is subject to strategic, and perhaps deceptive, stratagems, the suspect remains free to refuse to cooperate—to choose not to share thoughts and memories—without coercion, punishment, or prejudicial characterization of the suspect's silence at a potential trial.

These rights are human rights and so pertain no matter the suspect's nationality (United Nations 1948). Counter-terror operations may involve a state agent from state A interrogating a citizen of state B in state A, B, or C. One does not lose one's right to silence because a foreign agent finds one suspicious any more than one loses one's right if a domestic agent finds one suspicious (see Skerker forthcoming). Due-process rights are all the more relevant in this case because an agent is more likely to make mistakes when acting abroad than at home. If a suspected unprivileged irregular is captured

domestically in a liberal state, the individual ought to be treated as a criminal suspect. If agents act abroad, ideally they cooperate with local authorities to arrest a suspect and extradite the person or try them in local courts. Again, due-process rights also benefit the detaining power by reducing the likelihood that they capture an innocent person, leaving the real perpetrator at large. If one is fighting a large paramilitary organization in near-conventional battles, where combatants are more or less clearly identified due to their violent activities or uniforms, as is the case with ISIS, it may not be feasible or desirable to prosecute large numbers of detainees or to afford them access to lawyers. If, for practical reasons, the detaining power wishes to hold and process detainees as it would POWs in a conventional war—without access to lawyers and other due-process protections—it may. Yet, since this comes at the risk of wrongly holding innocent people, as with POWs, the detaining power must then forgo prosecution and *post-bellum* detention, because of the lack of due-process protections.

6. INTERROGATION TECHNIQUES

Having discussed the rights that counter-terror interrogators must observe, we can consider specific interrogation tactics with an eye to the practical elements of the security standard (proportionality, reliability, efficiency, and effectiveness). We will then consider how well these tactics do from a rights-respecting perspective. First, it is important to recognize that there are no special tactics suited for suspected unprivileged irregulars based on their political status, ideology, or chosen method of warfare. Not every tactic works for every person, but this disparity is rooted in differences of personality rather than martial status. Be it with POWs, domestic criminal suspects, positively identified, or suspected unprivileged irregulars—whether confessions or actionable Intelligence is sought—the key to eliciting information in interrogation is the development of rapport, a kind of respectful, working relationship where the interrogatee actually wants to talk to the interrogator (High-Value Detainee Interrogation Group 2016).

I have written extensively on this topic elsewhere (Hartwig et al. 2017, pp. 326–47).[2] In the interest of brevity, I will only focus on the scientifically validated, rapport-based techniques developed or validated in the last few decades. They all improve upon the *confession-based approach*, which dominated in many North American and European police departments in the second half of the twentieth century. With a confession-based approach, the interrogator cajoles, browbeats, deceives, and/or manipulates the suspect into confessing, overcoming their protestations of innocence with the interrogator's own version of events. The interrogator will often seek to maximize the suspect's sense of guilt—presenting confession as the only way to mitigate the

mental pressure—or minimize the seriousness of the crime and imply that the interrogator already knows enough to indict the suspect.

The first clinically validated novel approach is the *information-gathering model*. It sees the goal of interrogation as developing an accurate understanding of the event in question, rather than garnering a confession. The interrogator invites the suspect to present their own account of events, and engages in various exercises to help the suspect remember and present a full narrative. The interrogator then challenges inconsistencies in the suspect's account—inconsistencies that may be the result of the honest mistakes of an innocent suspect, or deliberate lies on the part of a guilty suspect.

Strategic interviewing takes into account that both interrogator and suspect (be the suspect innocent or guilty) are engaging in certain strategies in an interrogation. Strategic interviewing shares a basic approach with information-based methods in that the suspect is invited to narrate their account of the incident in question, but the interrogator may engage in certain behaviours that will prompt dishonest suspects to give away flagrant indications of their deception. The strategic interviewer attempts to prompt abnormal behavioural responses or inconsistent accounts of an incident (the interrogator's identification of which causes more stress for the suspect) by increasing the suspect's cognitive load, the amount of processes the suspect's brain must undertake at once. The interrogator might ask the suspect to maintain eye contact while narrating the account, describe events in reverse order, or draw a picture of the place the suspect claims to have visited. Liars have a hard time doing these things while maintaining a calm exterior since maintaining a fictional narrative is already cognitively taxing.

A type of strategic interviewing, the *Strategic Use of Evidence (SUE)* technique, similarly invites suspects to provide accounts of incidents without revealing how much information the interrogator already knows. The interrogator then selectively confronts the suspect with evidence contradicting their narrative. While people often innocently misremember events they have witnessed, narrate events out of order, or conflate sequences of events, truth-tellers and liars will typically respond in different manners when shown evidence contradicting their claims.

Another type of strategic interviewing, the *Scharff technique*, takes into account that a guilty suspect will seek to understand what information the interrogator already knows and then reveal as little new information as possible—without denying information the interrogator already knows. The Scharff interrogator accordingly seeks to give the suspect a false picture of what the interrogator already knows and wants to know. The interrogator presents a full picture of the suspect's biography, suspected actions, or network, making slight intentional errors, which the relaxed suspect usually hastens to correct. The interrogator then states suspected truths, which the suspect typi-

cally confirms or disconfirms. After a successful Scharff interview, the suspect does not know what information the interrogator sought and does not believe that the suspect gave the interrogator anything of value.

7. MORAL CONSIDERATION OF INTERROGATION TECHNIQUES

We will now consider how these techniques fare against the practical elements of the security standard (effectiveness, efficiency, proportionality, and reliability). Research on these novel techniques is ongoing. More research is clearly needed to permit more confident assessment of the following factors. Extant clinical and field research indicates that information-gathering approaches and strategic interviewing have a high degree of efficacy, reliability, and efficiency. These techniques do not seem to suffer from the high degree of unreliability associated with torture and the low-to-moderate unreliability of confession-based models (Hartwig et al. 2017, p. 331) since they do not rely on physically or psychologically pressing the suspect into confessing. Proportionality would compare the good produced with these techniques—accurate information—against the rights infringements associated with the techniques. Comparing the Intelligence yield of these techniques with the degree to which they infringe on suspects' rights shows that they fare better on a proportionality scale than do torture or confession-based approaches. These interrogation methods often elicit full, reliable, and accurate narratives from the suspect in relatively short order, and refrain from lying, threats, aggressive emotional manipulation, or violence (High-Value Detainee Interrogation Group 2016).

When it comes to the rights-respecting portion of the security standard, we can aver to John Rawls's famous 'veil of ignorance' thought experiment to model the sort of norm or tactic a generic person should endorse, bearing in mind that this person could be a member of the general public, an innocent or guilty suspect, or an interrogator. A way of testing a norm or tactic is to ask, 'Should I endorse this (practically efficacious) norm or tactic, or is it so onerous for the suspect and/or the interrogator that I would be unwilling to suffer it if I was wrongly suspected or if I was the interrogator?' A generic person ought to consider the extent to which a given interrogation approach respects the following rights relevant to interrogation: the suspect's autonomy in general, their right to privacy/silence, their right to honest dealing, their right to be respected, and their right to an attorney. Autonomy, in the sense of the capacity to determine what course of action one should take, relies, to an extent, on mental privacy and the assessment of true information. Mental privacy, as many conceive of it, is a kind of mental space where one can consider options and formulate ideas before sharing them with the world. This

'space' needs to be private and undisturbed for one to have ideas of one's own to act upon (Alfino and Mayes 2003; McCloskey 1980). The right to silence is an expression of a right to privacy in the moment when one is being pressed by others for personal information. The right to silence is comprehensive of the privilege against compelled self-incrimination insofar as deferring to the right to silence means the suspect will not be pressured, punished, or coerced into say anything, self-incriminatory or otherwise. The right to an attorney is the other due-process right relevant to the interrogation phase of an investigation, with the other rights relevant to the trial phase.

It will be helpful to consider what actions must *generally* be avoided to respect these rights as a prelude to considering how different novel interrogation techniques fare. Concerns about efficacy, which are moral concerns insofar as obtaining true confessions to criminal action or actionable Intelligence about imminent terrorist actions protect the rights of innocent people, might direct us to derogate from perfect respect for suspects' rights, as we will see. To respect someone's autonomy, one must refrain from taking actions that alter their understanding of what is true by emotionally manipulating them, lying to them, or deceiving them, or by forcing their choice by blackmailing them or using physical coercion against them. As mentioned above, autonomy is comprehensive of the right to honest dealing; lies and deception obviously risk violating a person's right to honest dealing.

One must treat a person politely, and in a way that honours their dignity, in order to meet their right to be respected. One respects a person's right to an attorney by telling the individual that they have that right, making it possible to contact one, arranging a meeting with an impartial attorney if the suspect cannot find or afford one, allowing the attorney to sit in on the entire interrogation, and not penalizing the suspect in any way for invoking the right to an attorney. To respect a suspect's right to privacy and silence, one ought to admonish the individual that they do not have to cooperate with interrogators and will not be penalized for silence. The suspect ought not to be pressed with personal questions, particularly if they object; the suspect ought not to be tricked into revealing what they wish to conceal; and the suspect ought not to be penalized or threatened with any sort of penalty if they refuse to cooperate.

The information-gathering approach scores very well with respect to the rights element of the security standard because of the steps it takes to establish rapport and maximize a suspect's ability to remember information. The interrogator explains the investigative process to the suspect and the suspect's rights, treating the individual at all times with respect and eschewing any kind of deceptive or aggressive behaviour. Since the technique relies on the suspect presenting a narrative, even a fictional one, the skilled interrogator will want to project a solicitous air that encourages the suspect to share with the interrogator. As in all interrogations, a lawyer might well counsel the client's silence,

but there is nothing untoward about the information-gathering approach that a lawyer would object to per se. On the subject of lawyers—since some laypeople assume that terrorists 'lawyering up' reduces the efficacy of interrogations—it should be noted that, since US detectives have been required to issue the Miranda warning advising suspects of their rights to an attorney and to silence, most suspects have nonetheless waived their rights and spoken to detectives (Ward 2016). It is likely that many innocent suspects waive their rights to silence because they want to cooperate with detectives to clear up the apparent misunderstanding that led them to be considered suspicious—without incurring legal fees—and that inexperienced guilty suspects fear that refusing cooperation will make them appear guilty and prolong both interrogation and investigation. An advantage that the novel interrogation techniques have over confession-based models is that they do not open with an accusation of guilt. It follows that interrogators eschewing the confession-based model can have an easier time gaining the trust of suspects, which in turn might make requesting a lawyer seem unnecessary. It seems even less likely in counter-terror contexts that suspects from developing countries with immature legal institutions would think it important to have an attorney, even if they were offered one.

Regarding the rights to privacy and silence, the information-gathering approach does not rely on trickery, manipulation, or threats. Overt manipulation and threats might inhibit rapport-building. To be sure, there is a privacy-infringing aspect to the information-gathering approach in that an authority figure invites the suspect to give a narrative of a certain event. This behaviour would be generally inappropriate for a private citizen to do to a stranger. Yet, this degree of intrusion would be present with any interrogation style.

The SUE approach similarly relies on rapport between interrogator and suspect and so eschews disrespectful or aggressive behaviour, including lies, threats, blackmail, or any kind of physical coercion. The interrogator wants the suspect to feel comfortable and voluble; this state is encouraged by the interrogator explaining to the suspect the scope of their rights and the interrogation process. An admonition about the suspect's right to silence is of a piece with this solicitous approach. Holding a suspect's silence against the suspect (say, through penalty of fines, incarceration, or a judge's negative characterization of the suspect's silence to an eventual jury) would ill serve the SUE approach since this threat would function against the development of rapport. As with the information-gathering approach, a lawyer would likely counsel the client to refuse all cooperation, but the SUE approach does not rely on accusations of guilt or aggressive behaviour that might prompt the suspect to immediately request a lawyer or refuse cooperation.

The SUE approach infringes on the suspect's autonomy and privacy more than the information-gathering approach. The SUE interrogator passively

deceives the suspect by failing to initially reveal the incriminating evidence held against the suspect. The interrogator engages in strategic behaviour, inviting the suspect to commit to false accounts of an event and then creating additional psychological stress by revealing the evidence contradicting that account. The interrogator also invites the suspect to engage in certain behaviours, like drawing a picture of the room in which an event supposedly occurred, which the interrogator knows will cause stress for a deceitful suspect. The suspect's flustered reaction to this stress and to the interrogator's gentle confrontations regarding the deception may cause the suspect to reveal more information than originally planned. The suspect does not have a right to withhold criminal information, but may also reveal personal non-criminal information in the course of these protestations and justifications.

The Scharff technique also avoids disrespectful, threatening, or coercive behaviour. As it is the most reliant on rapport of the three novel techniques, the Scharff interrogator wants to avoid any behaviour that would make the suspect feel defensive. Not only does the admonition about silence not hinder Scharff, a form of the admonition is part of the standard Scharff preamble. The interrogator tells the suspect that the former already knows all about the latter and doubts that there is anything relevant that the suspect can add. The interrogator then narrates all that is known, trying to communicate to the suspect that there is no point in remaining silent because the interrogator already knows everything. The same dynamic would likely make hiring a lawyer seem gratuitous. The Scharff interrogator would likely be the least perturbed among other types of interrogators in the event that a lawyer is hired because the interrogator does not initially rely on narratives from the suspect.

As with the SUE technique, the Scharff technique can infringe on the suspect's autonomy and rights to honest dealing and privacy. The interrogator deceives the suspect regarding what the interrogator already knows, and more so than the SUE interrogator, projects a perhaps false 'I know all' effect. More so than the other techniques, Scharff also requires the interrogator to project an urbane, solicitous, and friendly demeanour. While even the information-gathering approach requires the interrogator to project a courteous and respectful attitude to a person the interrogator may actually despise, the successful Scharff interrogator will make the suspect feel like they just spent a delightful afternoon with someone, who, but for the vagaries of citizenship, might have been an esteemed colleague. Lulled by the interrogator's friendliness and seemingly comprehensive knowledge of the suspect, the suspect may divulge more private, non-criminal information than they would otherwise have chosen to share.

In sum, the information-gathering approach is most respectful of rights among the novel techniques. SUE arguably scores lower with regard to the rights-respecting element of the security standard than Scharff. While both

deceive through omission and treat the suspect strategically, SUE is likely more onerous for the suspect since it involves confrontation over contradictions in the suspect's story and the presentation of incriminating evidence. The general unpleasantness of an interrogation approach is relevant to the security standard because we model professional norms by asking what people would tolerate as the cost of capturing unprivileged irregulars in the event that they were wrongly suspected and subjected to interrogation. All this said, there do not seem to be huge disparities between the novel techniques with respect to their rights-respecting abilities.

The security standard prefers norms and tactics that are the most rights-respecting out of those that are the most practically efficacious. Thus, if all three novel techniques are equally efficacious, interrogators ought to employ the information-gathering approach. One of the other two techniques would instead be the preferable one if further research indicated significant deficiencies in the information-gathering approach's relative efficacy.

NOTES

1. This chapter draws on arguments I develop in greater length in Skerker (2010) and Skerker (2020).
2. See this article for a bibliography on interrogation techniques.

REFERENCES

Alfino, Mark and G. Randolph Mayes (2003), 'Reconstructing the Right to Privacy', *Social Theory and Practice* 29(1), 10.

Arrigo, Jean Maria (2004), 'A Utilitarian Argument against the Torture Interrogation of Terrorists', *Science and Engineering Ethics* 10(3), 543–72.

Camenisch, Paul (1983), *Grounding Professional Ethics in a Pluralistic Society*, New York: Haven.

Edmonds, Bill Russell (2015), *God Is Not Here*, Berkeley, CA: Pegasus Books.

Hartwig, Maria, Timothy Luke, and Michael Skerker (2017), 'Ethical Perspectives on Interrogation: An Analysis of Contemporary Techniques', in Jonathan Jacobs and Jonathan Jackson, eds., *The Routledge Handbook of Criminal Justice Ethics*, New York: Routledge, pp. 326–47.

High-Value Detainee Interrogation Group (2016), HIG Interrogation Best Practices Report, accessed 26 March 2021, at https://www.fbi.gov/file-repository/hig-report-august-2016.pdf/view.

Lagouranis, Dan (2008), *Fear Up Harsh*, London: Dutton Caliber.

McCloskey, Henry J. (1980), 'Privacy and the Right to Privacy', *Philosophy* 55(211), 21.

Miller, Seumas (2010), *The Moral Foundations of Social Institutions*, Cambridge: Cambridge University Press.

Phillips, Joshua (2012), *None of Us Were like This Before*, London: Verso.

Skerker, Michael (2010), *An Ethics of Interrogation*, Chicago, IL: The University of Chicago Press.

Skerker, Michael (2011), 'The Rights of Irregular Combatants', *International Journal of Intelligence Ethics* 2(1), 35–49.
Skerker, Michael (2019), 'What Can Be Asked of Interrogators?', in Steven J. Barela, Mark Fallon, Gloria Gaggioli, and Jens David Ohlin, eds., *Interrogation and Torture*, Oxford: Oxford University Press, pp. 253–78.
Skerker, Michael (2020), *The Moral Status of Combatants*, London: Routledge.
Skerker, Michael (forthcoming), 'The Rights of Foreign Intelligence Targets', in Seumas Miller, Adam Henschke, and Jonas Feltes, eds., *Intelligence Ethics*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.
United Nations (1948), *Universal Declaration of Human Rights*, accessed 1 April 2021, at https://www.un.org/en/about-us/universal-declaration-of-human-rights.
Walzer, Michael (2015), *Just and Unjust Wars*, 5th ed., New York: Basic Books.
Ward, Stephanie Francis (2016, May 23), 'How Well Do People Actually Know Their Miranda Rights?', interview with Russell L. Covey, American Bar Association podcast, accessed 26 March 2021, at https://www.abajournal.com/news/article/podcast_monthly_episode_75.

7. Preventive detention of terrorists

Seumas Miller

1. INTRODUCTION

In this chapter, our concern is with the preventive detention of terrorists and, specifically, whether preventive detention of terrorists might be morally justified.[1] Accordingly, we need to have serviceable accounts of preventive detention and of terrorism. Regarding terrorism, the following definition will be relied upon (see Chapter 2):

Terrorism is a political or military strategy that:

1. Consists of state or non-state actors deliberately perpetrating acts of violence aimed at (directly or indirectly) seriously physically harming, and typically killing, innocent civilians;
2. Is a means of terrorizing the members of some social, economic, political, ethnic or other group to achieve a political purpose;
3. Relies on the violence receiving a degree of publicity, at least to the extent necessary to engender widespread fear in the target group.

What of preventive detention? Preventive detention is a portmanteau notion used to refer to various forms of detention, including prisoner-of-war camps; quarantine during epidemics; detention of illegal immigrants; extensions of imprisonment terms beyond their initial sentences for 'dangerous' offenders, such as paedophiles; and short-term detention without charge of those suspected of intending to conduct an imminent terrorist attack. Here, we are primarily concerned with the imprisonment of members of terrorist organizations for the principal purpose of preventing future terrorist attacks by that organization (whether those future attacks be imminent or not) and, therefore, for the purpose of preventing harm by the persons imprisoned but also for the purpose (by means of interrogation) of collecting information from them regarding the organization's planned attacks, membership and so on. Accordingly, as with all categories of preventive detention, the purpose is forward-looking; the purpose is not, for instance, punitive and backward-looking, as in the case of punishing terrorists for past offences. Second, a fundamental purpose of pre-

ventive detention is to prevent harm to the community, including the murder of members of the community. Thirdly, preventive detention, as it is used here, does not refer to the related phenomenon of control orders. The latter do not involve imprisonment but rather the placing of various restrictions on the movements (for example, confinement to a certain address), communications (for example, prohibition on using a phone or email) and so on of known or suspected offenders, including, for instance, so-called returning foreign fighters, for example, citizens of the UK and elsewhere who were known or suspected of having travelled to Iraq or Syria to fight for the Islamic State but who have since returned home.[2]

2. TERRORISM AND PREVENTIVE DETENTION

According to our definition, terrorism is a violent means in the service of a political end. Moreover, the violence is typically directed at persons who are considered to be innocent, for example, civilians as opposed to combatants. Accordingly, terrorism is, or ought to be, a crime. In the case of state terrorism, in particular, it is sometimes not a crime, at least in the state perpetrating the acts of terror in question.

Combating the crime of terrorism involves particular difficulties not typically present in combating other crimes. One difficulty revolves around the status of terrorists. According to one view, they are simply criminals to be dealt with by police officers operating within a law-enforcement framework. But terrorists often insist that they are political actors fighting a war, that is, that they are military combatants. Indeed, in some instances, terrorists are clearly de facto (if not *de jure*) combatants (even if they are illegal combatants, as the United States declared members of al-Qaeda to be; see Blum 2008, Ch. 3), such as members of the Islamic State engaged in large part in conventional warfare in Syria and Iraq. However, during wartime, civilians are a separate category from combatants, and the rules of engagement with enemy combatants do not pertain to non-combatants. Moreover, terrorist-combatants are not simply combatants since, qua terrorists, they are criminals; terrorism is, after all, a crime in most jurisdictions and under international criminal law. Moreover, this is the case even if they are deemed to be unlawful combatants on some ground other than the fact that they engage in terrorism, for example, they do not wear uniforms and bear arms openly. Further, terrorists who are de facto combatants and who engage in, for instance, the murder of civilians in theatres of war are guilty of *war* crimes. Moreover, in all this there are complications arising from differences, firstly, between terrorists who are citizens or residents of the state under attack from the terrorist organization in question and terrorists who are foreigners and, secondly, between terrorist-combatants

captured on the battlefield and terrorists arrested in well-ordered jurisdictions outside areas of active hostilities.[3]

Terrorists pose problems greater than other dangerous criminals by virtue of the fact that they typically constitute an organized group that deliberately targets large numbers of people and does so indiscriminately – that is, members of the public at large are the targets. Given the danger to ordinary citizens posed by terrorists and, in particular, the need to prevent terrorist acts rather than merely react to them once they have been committed, the preventive detention of terrorists is an attractive option for governments. However, the preventive detention of suspects and the detention of suspects for prolonged periods without their being charged and tried are infringements, even if not violations, of the human right to freedom of action.[4]

Indeed, the cornerstone of liberal democracy is individual freedom and, aside from freedom of thought and speech, the most fundamental freedom, or set of freedoms, is freedom of action. Freedom of action includes freedom of bodily movement, freedom to associate and form relationships with others, freedom to buy and sell, freedom to plan and implement projects, including one's career, and so on. It is self-evident that detention, and especially long-term imprisonment, strike at the very heart of individual freedom. For this reason, imprisonment ought to be reserved only for serious crimes and in circumstances in which the suspect is guilty beyond reasonable doubt, or so it would seem. Thus, detention for prolonged periods without trial is morally unacceptable. Faced with these kinds of individual rights-based arguments, a tendency has developed on the part of governments to invoke the notion of trade-offs, and a balance between individual rights, on the one hand, and security considerations, on the other; this is especially the case in relation to anti-terrorist legislation.[5]

Here there are two crucial issues. The first regards whether or not there is in fact a need for a trade-off and, specifically, a trading down of particular individual rights. Arguably, privacy can be traded down to a significant degree, but freedom of action cannot (Kleinig et al. 2011). Or, perhaps we can increase security by spending more money (and time) on, for example, airport security, surveillance of at-risk installations and border controls without any significant diminution of existing privacy rights or existing rights to freedom. Secondly, in so far as there is a need for balancing and to trade off, what is to be put on the scales, and what is to be traded off against what?

With respect to one side of the scale, what proponents have in mind is perhaps clear enough; individual freedom is on the scales and is to be traded down. However, it is the other side of the scales that is unclear. Notions of national security or community safety are far too general and vague to be helpful here. There is a need for more precise and differentiated notions. Indeed, as far as the notion of community safety is concerned, this presumably

largely consists in the human rights to life and other aspects of personal security; so the other side of the scales consists in an individual right, after all, viz. the right to personal security. As is often the case, balancing rights to freedom and rights to personal security – if this is what has to be done – is a complex matter; sometimes the latter will trump the former – for example, searching luggage for bombs at airport security points – and there are contexts in which the former will trump the latter – for example, British soldiers going to war against Hitler's Nazi forces.

However, it is by no means clear that there is a need for a trade-off between fundamental rights to individual freedom and rights to personal security in well-ordered, liberal democratic states at peace. For one thing, security consists in large part in the provision of the conditions for the exercise of individual freedom. National security and law and order in liberal democratic states, as has been argued elsewhere (Miller and Blackler 2016, Ch. 1), largely consist in, or are heavily dependent on, respect for human and other moral rights, especially rights to personal security and property rights. Without respect for personal security and respect for property rights, there is no law and order in a liberal democracy and, therefore, the exercise of individual liberty is difficult, if not impossible. For another thing, the trade-off can be, and ought to be, a trade-off between the rights of offenders and suspected offenders on the one hand, and the rights of innocent people on the other. It is not as if what are to be traded down are the rights to, say, life and liberty of innocent civilians. The proposition is not that police and other security personnel ought to be empowered to shoot to kill, or indefinitely detain, *innocent* people in order to protect the rights of other innocent people.

While politicians in liberal democracies frequently frame the issue of preventive detention in terms of the trade-off between individual rights and national security, those who oppose preventive detention focus on the human rights of those preventively detained who are merely suspected, but not convicted, of terrorism and, therefore, might not in fact have perpetrated any act of terrorism (Blum 2008; Webber 2016). However, the existence of terrorist-combatants is problematic for this rights-based law-enforcement perspective. Let us now turn directly to these issues.

3. PREVENTIVE DETENTION, TERRORISTS AS CRIMINALS AND TERRORIST-COMBATANTS

Suspects are, by definition, not identical to those who have been tried and found guilty of a crime. Thus, unlike those who have been tried and found guilty, suspects continue to be presumed to be innocent and, as a consequence, cannot be, or ought not be, detained for lengthy periods, or otherwise subjected to restrictions or harms. Rather, suspects who are arrested must surely either be

charged and brought to trial expeditiously, or must be released (perhaps after a restricted period of interrogation). Moreover, suspects who are subjected to detention and interrogation ought to be afforded appropriate rights to protection, for example, the right to an attorney (Regan and White 2021, Ch. 1).

This is not to say that there might not be a need to calibrate, for example, periods of detention without trial in the context of changing circumstances, including the current threat of terrorism in the US, UK, France and elsewhere. Thus it may be that terrorist suspects ought to be able to be detained for weeks rather than days in the context of, for instance, the need to extract evidence from encrypted communications on seized computers. But such calibration must not be assimilated to a circumstance in which a terrorist suspect can be detained indefinitely without trial (including by the device of ongoing renewal of a detention order) as has been the case in some jurisdictions, for example, the United States' Guantanamo Bay prison camp (Blum 2008, Ch. 2).

Preventive detention is a controversial counter-terrorist measure that certainly infringes, and perhaps violates, the individual moral right to freedom. On the one hand, it is claimed by some (for example, human rights advocates and organizations, such as Human Rights Watch; see Fathi 2009) to be a human rights violation since it involves imprisonment of suspected terrorists who have not been tried for terrorism and found to be guilty beyond reasonable doubt in accordance with due process of law. On the other hand, others (for example, members of the former Bush administration in the US) have argued that at least some terrorist are combatants, although unlawful combatants, and can be subjected to preventive detention as combatants and perhaps (unlike ordinary prisoners of war) subjected to interrogation by virtue of not having the rights of lawful combatants undergoing detention (Blum 2008, Ch. 2).

Two conceptually separable moral justifications for preventive detention are embedded in the above-mentioned controversy. Firstly, there is the justification based on terrorism understood as a serious crime. Secondly, there is the justification based on terrorists as dangerous, irrespective of whether or not they are morally (or legally) culpable or even morally (or legally) responsible for their dangerousness. In relation to this second justification, consider enemy combatants or persons held in quarantine.

Regarding terrorists as criminals, the preventive detention of terrorists is morally problematic in that, at least in principle, it does not necessarily pertain to those suspected of a past or present crime – let alone tried and convicted of a crime – but to those suspected of being likely to commit a *future* crime; that is, persons are to be detained, notwithstanding the fact that the crime for which they are being detained has not been committed and is not in the process of being committed. Here it is important to distinguish between: (a) someone suspected of having already committed a crime – this first crime is in the present – as a precursor to committing a second crime in the future – for

example, conspiring in the present to commit a murder in the future; and (b) someone who is not suspected of any present (or past) crime, but only of being likely to commit a future crime – for example, someone who is not suspected of any past or present crime, such as the crime of conspiracy to murder, but who is, nevertheless, believed to be likely to commit a murder in the future. At least in principle, preventive detention might pertain only to a person in the situation described in (b), and not to a person in the situation described in (a). As such, preventive detention infringes the basic moral principle that a person should not be detained, or otherwise penalized, for a crime that they are known not to have committed or to be in the process of committing. Accordingly, so the argument runs, preventive detention cannot be morally justified.

What of the idea that terrorists are dangerous (irrespective of their moral or legal culpability for their dangerousness)? Thus understood, preventive detention might be morally justified by analogy with enemy combatants or those held in quarantine, depending on the quantum of innocent lives terrorists or terrorist organizations put at risk. Naturally, this justification has its limits. For instance, preventive detention might not be necessary if terrorist attacks are able to be thwarted utilizing less morally questionable means; and preventive detention might be disproportionate (and perhaps counter-productive) given (say) if the practice in the context in question required the detention of thousands of suspected terrorists over many decades, and yet only a small number of innocent lives would be put at risk if preventive detention was eschewed in favour of less morally questionable means.

Given that terrorists are both *criminals* guilty of past or present serious moral wrongdoing and *highly dangerous* persons likely to perpetrate future acts of murder – and, importantly, morally responsible for their dangerousness, unlike those held in quarantine, for instance – it seems that an adequate justification for preventive detention would need to help itself to both of these moral considerations. In doing so, it might not be relying on the disjunctive view that a terrorist is *either* a criminal *or* a combatant, but rather on the conjunctive view that, at least in the case of the members of terrorist organizations who can engage in armed conflict, such as al-Qaeda and the Islamic State, a terrorist is *both* a criminal *and* a combatant (albeit an unlawful combatant). Before addressing this issue in detail, two points should be made in passing.

Firstly, whether or not the preventive detention of terrorists in the form of long-term imprisonment is morally justified, the preventive detention for limited periods in some emergency situations is surely justified. For example, in the context of ongoing, large-scale, caste-based and communal violence of the sort experienced in Bihar and Gujarat in India in recent decades (Miller et al. 2008), preventive detention for limited periods of persons highly likely to incite mass crowds to violence might be morally justified. However, this is a moral justification for the preventive detention of select individuals for

a limited period and only in the context of a well-founded, and lawfully decreed, state of emergency.

The second point, in relation to the issue mentioned above of trading down of the rights of, especially, terrorist suspects, is this one. One illicit way in which the scales on one side (the security side) are being weighed down with a consequent trading down of the rights of suspects on the other side, is by the broadening of the scope of anti-terrorist legislation so as to embrace not simply actual specific acts of terrorism or actual membership in terrorist organizations, but also *threatened* acts of terrorism and the consequences of actual acts of terrorism in terms of the *fear* that they might produce. In some jurisdictions (Bottomley and Bronitt 2006, p. 402) terrorism includes the (possibly indirect and distant) *threat* of bombings and like actions, and therefore brings with it actions that have the potential to cause harm, for example, undertaking terrorist training; moreover, some anti-terrorism laws also focus on the motivation to intimidate and therefore bring into play the intentional causing of the *fear* of harm, as opposed to harm itself. There are other ways of widening laws against terrorism – for example, associating with a terrorist – and new crimes (or the resuscitation of ones in disuse) – for example, sedition. Here, as elsewhere, there is a need to analyse each of these elements on a piecemeal basis. Undergoing terrorist training, for example, manifests a high degree of culpability and, in the context of an increasing terrorist threat, warrants severe penalties. On the other hand, whether or not an action intentionally or otherwise caused fear is arguably so indeterminate a matter as to lead to abuse in the application of any laws enacted to eliminate or reduce such fear-causing actions.

4. PREVENTIVE DETENTION OF TERRORIST-COMBATANTS

Thus far we have seen that, at least some terrorists, such as members of al-Qaeda and the Islamic State, are both terrorists and combatants, that is, terrorist-combatants. Let us now consider preventive detention in relation to terrorist-combatants.

The preventive detention of terrorist-combatants is evidently justified by the moral principles, if not the laws, governing the conduct of war.[6] Since terrorist-combatants are combatants, and it is legally and morally justifiable to incarcerate captured combatants until the cessation of hostilities to prevent them from resuming the fight, by parity of reasoning, it is morally justifiable to preventively detain terrorist-combatants until the cessation of hostilities.

Moreover, since terrorist-combatants are combatants, it is legally and morally justifiable for combatants in an opposing force to kill them. One salient moral principle here is the one already mentioned, namely, that com-

batants are not only dangerous but also morally responsible for their dangerousness.[7] In this respect, combatants are unlike, for instance, infected persons in quarantine. A second salient moral principle is that of a *standing intention*. Combatants, by virtue of their occupancy of a role in an armed force engaged in armed conflict, are reasonably presumed to have a standing intention to kill combatants in the opposing force (and will do so unless the latter intervene to protect themselves by killing the former). Roughly speaking, standing intentions activate immediate intentions, which in turn cause actions. However, one can have a standing intention without having a relevant immediate intention – for example, a combatant who is eating lunch in a secure building – and one can have an immediate intention without having a standing intention – for example, a husband who intentionally kills his adulterous wife in a fit of anger but without any premeditation. The difference between combatants and terrorist-combatants is that the latter have a standing intention not only to kill enemy combatants but also to kill innocent civilians, should they be ordered to do so. Hence terrorist-combatants are unlawful combatants and, indeed, morally culpable combatants.

It is important to note that the dangerousness – understood as being comprised in large part of a standing intention and an ability to harm – of enemy combatants and, therefore, of terrorist-combatants involves the interdependence of standing intentions and a jointly held ability to harm. An individual combatant only has a standing intention to harm if they are a member of an organization in which their fellow members also have a standing intention to harm – and a standing intention to harm in the service of the same shared end, for example, to win a war. Accordingly, if all but one of the members of an army lay down their arms when the cessation of hostilities is declared, then the remaining one will typically do so. Moreover, the harm that an individual combatant can cause acting on their own is typically quite limited relative to what the armed force as a whole can cause. Further, the individual combatant does not act on their own as an individual but rather qua member of an organization. They act under orders from others in accordance with a strategy devised by others, and the tasks they are set are typically joint tasks, for example, take and hold a strategically important hill currently occupied by the enemy. This raises the issue of the *collective moral responsibility* – understood as joint moral responsibility (Miller 2006, pp. 176–93) – of combatants for harms resulting from their joint action as opposed to harm for which a single individual is solely morally responsible (see Chapter 3). Thus, a single combatant might be individually morally responsible for killing the enemy combatant they shot dead, but also morally responsible – jointly with others – for defeating the enemy platoon of which that enemy combatant was a member.

It is also important to note that the preventive detention of enemy combatants is *not* typically indefinite, even if it has lasted for an extended period

of time. Rather, prisoners are to be released at a definite end point, namely, the cessation of hostilities. By parity of reasoning, terrorist-combatants qua combatants ought to be released at the cessation of hostilities: presumably, a definite end point, for example, such as obtained in the case of the Irish Republican Army's (IRA) cessation of its terrorist campaign in the UK. It might be suggested that, unlike wars, terrorist activities go on indefinitely. As the IRA's terrorist campaign and numerous terrorist campaigns in anti-colonial struggles demonstrate, this is not necessarily or even typically the case. In any case, for our purposes here we need to distinguish terrorists engaged in armed conflict, that is, military-style campaigns, from sporadic bombings of civilian targets or armed 'marauders' in well-ordered jurisdictions (such as occurred in Paris in November 2015 at the Bataclan and other locations). It is the former and not the latter that is in question at this stage in the argument.

In the above discussion it has been argued that terrorist-combatants are combatants and, therefore, can be preventively detained until the cessation of hostilities. However, terrorists are unlike lawful combatants in two respects. Firstly, terrorist-combatants kill innocent civilians. Secondly, terrorist organizations defeated on the battlefield can continue to engage in acts of terrorism in well-ordered jurisdictions. Thus citizens of, say, the UK travel overseas and join a terrorist organization, such as the Islamic State; function as terrorist-combatants in a theatre of war, such as Iraq or Syria; and then return to the UK to carry out terrorist attacks in the UK as members of the Islamic State. These are the so-called foreign fighters that domestic security agencies worry so much about – home-grown terrorists with battlefield experience (Hoffman 2019). However, former terrorist-combatants who return to their country of origin to carry out sporadic terrorist attacks on ordinary civilians in their well-ordered home jurisdictions are not thereby engaging in armed conflict. Therefore, they are not terrorist-combatants by virtue of carrying out these terrorists attacks in their well-ordered home jurisdictions, as opposed to doing so on the battlefield.

Importantly, terrorists, including terrorists who are not combatants, are collectively, that is, jointly, morally responsible for the murders committed by the terrorist organization of which they are functionally integrated members, in addition to being individually morally responsible for whatever contributory actions they perform. Thus, the member of a terrorist organization who trains other members to, say, use explosives to murder people is doubly morally culpable. First, they are *fully individually* morally responsible for providing those they train with the means to murder. Second, qua functionally integrated member of the terrorist organization, they are, *jointly with the other members*, morally responsible for the murders performed by multiple members of the organization, albeit these murders are not actually performed by them, and their responsibility may only be partial (Miller 2010, Chs. 1 and 2).

This 'double' moral culpability has an important potential legal implication, namely, that the terrorist in question can reasonably be held criminally liable not only for the crime of training members of a terrorist organization, but also for the crime of murder or, at least, complicity in murder (Blum 2008) via their functionally integrated membership in a terrorist organization whose core business is murder.

There is a further important implication of being a functionally integrated member of a terrorist organization that is central to our concern with preventive detention. Being a functionally integrated member of a terrorist organization entails that the terrorist member in question has a standing intention – and a jointly held ability – to commit murder or to assist others to commit murder. Therefore, the presumption ought to be that such a person will commit murder or assist others to do so unless prevented from doing so.[8] Accordingly, there ought to be a presumption that a convicted terrorist who is imprisoned for their terrorist crimes will commit murder or assist others to do so, if they are released. Naturally, this presumption can be overridden and the detainee in question is entitled to periodic reviews to determine whether the presumption should be overridden. For instance, the presumption would be overridden if the terrorist organization in question abandoned its policy and practice of murder, or if the erstwhile terrorist demonstrates that they now reject terrorism. However, if this presumption is not overridden, then the terrorist in question can reasonably be preventively detained and for the same reason that enemy combatants are held as prisoners of war, namely, that by virtue of their functionally integrated organizational membership they have a standing intention to kill and will do so unless prevented from doing so by incarceration.

It has been argued that the members of terrorist organizations engaged in armed conflict, such as al-Qaeda and the Islamic State, can be presumed to have a standing joint intention to kill innocent civilians – whether in theatres or war or in well-ordered jurisdictions – and that this justifies their incarceration until the cessation of hostilities (irrespective of whether they were captured in a theatre of war or not). However, their individual moral culpability, taken in conjunction with the magnitude of the threat of the terrorist organization to which they belong, raises the question: Should some of the rights enjoyed by lawful combatants, on the one hand, and by ordinary criminals, on the other, be curtailed? For instance, lawful combatants do not have to provide intelligence to their captors (other than name, rank and serial number). Again, ordinary criminals typically have a right to silence.[9] Arguably, it should be permissible to interrogate members of such terrorist organizations, and their right to silence should be abrogated. However, interrogatory torture should not be permitted, nor should the right not to self-incriminate be abrogated; or, at least, terrorists should enjoy immunity from prosecution if they self-incriminate.

5. CONCLUSION

It was argued above, firstly, that since terrorist-combatants are de facto combatants, it is morally justifiable for them to be preventively detained in the manner of prisoners of war, that is, until the cessation of hostilities. Secondly, it was argued that even terrorists who are not terrorist-combatants can justifiably be preventively detained if it is established that they are functionally integrated members of a terrorist organization. This is most likely to be established by recourse to actions undertaken on behalf of the terrorist organization that are crimes in their own right. These might be lesser crimes than murder or even attempted murder – for example, the ongoing provision of training, recruitment or finance.

However, there is an important remaining question as to the standard of proof required to establish that a person is a functionally integrated member of a terrorist organization. Arguably, the standard of proof should be that of 'beyond reasonable doubt', at least for long-term detention. Presumably, the standard of 'beyond reasonable doubt' in relation to the conviction of a person for the crime of membership in a terrorist organization would be met if the person in question was convicted of, for instance, training terrorist members of that organization, and the standard of proof met in relation to this lesser crime was 'beyond reasonable doubt'. On the other hand, the standard of 'on the balance of probabilities' might be sufficient for short-term detention – for example, periods up to six months (assuming hostilities continue) – in order to avert the near-term harm of terrorist attacks currently being planned. Moreover, in such cases, the short period of detention might be followed by a period in which restrictions are placed on the movements, communications and so on of the person in question (that is, some form of control orders).

NOTES

1. For an argument justifying the preventive detention of terrorists that is somewhat different to the one presented here, see Scheid (2010). For discussion of Scheid, see Landesman (2011). For replies to Landesman and others, see Scheid (2011). For a more recent discussion of Scheid, see Miller (2018).
2. See Webber (2016), Chapters 5 and 6, for discussions of control orders.
3. There are also grey areas that are neither battlefields nor well-ordered jurisdictions, for example, the FATA in Pakistan. See Miller (2009), Chapters 4 and 5, for discussion of the significance of this threefold distinction.
4. See, for instance, Webber (2016) for a detailed treatment of the legal principles and issues from a human rights perspective.
5. See, for example, what Philip Ruddock, the former Australian Attorney General, had to say about this. He is quoted in Bottomley and Bronitt (2006, p. 412).
6. See Miller (2018, pp. 122–40).

7. Douglas Husak (2011) emphasizes that many dangerous criminals are responsible for their dangerousness.
8. Note that those who assist terrorists qua terrorists can be functionally integrated members of a terrorist organization. The notion of a functionally integrated member of an organization is essentially that of a role defined in terms of tasks and an occupant of that role; the occupant pursues the collective ends of the organization and does so by virtue of occupancy of the role. See Miller (2010), Chapters 1 and 2.
9. On interrogatory torture, see Skerker (2021), which is Chapter 6 in this volume.

REFERENCES

Blum, Stephanie (2008), *The Necessary Evil of Preventive Detention in the War on Terror*, New York: Cambria Press.

Bottomley, Stephen and Simon Bronitt (2006), *Law in Context*, 3rd ed., Sydney: Federation Press.

Fathi, David (2009), 'Dangers of a Preventive Detention Law', Human Rights Watch, accessed at https://www.hrw.org/news/2009/01/03/dangers-preventive-detention-law.

Hoffman, Bruce (2019), 'Understanding the Evolving Terrorist Threat Landscape', presentation given at NSI's Impact Conference, National Security Institute, accessed at https://nsi.org/Impact19Presentations/B.Hoffman.pdf.

Husak, Douglas (2011), 'Lifting the Cloak: Preventive Detention as Punishment', *San Diego Law Review* 48(4), 1173.

Kleinig, John, Peter Mameli, Seumas Miller, Douglas Salane, and Adina Schwartz (2011), *Security and Privacy: Global Standards for Ethical Identity Management in Contemporary Liberal Democratic States*, Canberra: ANU Press.

Landesman, Bruce (2011), 'Scheid's Dilemma', *Criminal Justice Ethics* 30(1), 98–105.

Miller, Seumas (2006), 'Collective Moral Responsibility: An Individualist Account', *Midwest Studies in Philosophy* 30(1), 176–93.

Miller, Seumas (2009), *Terrorism and Counter-Terrorism: Ethics and Liberal Democracy*, Oxford: Blackwell.

Miller, Seumas (2010), *The Moral Foundations of Social Institutions: A Study in Applied Philosophy*, New York: Cambridge University Press.

Miller, Seumas (2018), 'The Moral Justification for Preventive Detention', *Criminal Justice Ethics* 37(2), 122–40.

Miller, Seumas and John Blackler (2016), *Ethical Issues in Policing*, London: Routledge.

Miller, Seumas, Sankar Sen, Prakash Mishra, and John Blackler (2008), *Ethical Issues in Policing in India*, Hyderabad: National Police Academy.

Regan, Mitt and Alexandra L. White (2021), 'Preventive Criminal Law', in Seumas Miller, Adam Henschke, and Jonas Feltes, eds., *Counter-Terrorism: The Ethical Issues*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Scheid, Don (2010), 'Indefinite Detention of Mega-Terrorists in the War on Terror', *Criminal Justice Ethics* 29(1), 1–28.

Scheid, Don (2011), 'Replies to Commentaries', *Criminal Justice Ethics* 30(1), 111.

Skerker, Michael (2021), 'Interrogation Ethics in Counter-Terrorist Operations', in Seumas Miller, Adam Henschke, and Jonas Feltes, eds., *Counter-Terrorism:*

The Ethical Issues, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.
Webber, Diane (2016), *Preventive Detention of Terror Suspects*, London: Routledge.

8. Use of stings in counter-terrorism: entrapment and ethics

Seumas Miller

1. INTRODUCTION

This chapter provides an ethical analysis of the use of stings or traps in police counter-terrorism operations in liberal democracies and, specifically, in the US. The conclusion reached is that such operations are morally justified under certain conditions, such as that they are effective, there are no less-intrusive methods available, and the inducements offered are ones that the targets of the sting could reasonably be expected to resist. An important novel condition that is proposed is that the suspect has a standing intention (and/or an immediate intention) to commit a serious terrorist offence – for example, they occupy a role in a terrorist organization as a bomb-maker, recruiter or trainer – or has otherwise expressed the standing and/or immediate intention to commit a serious terrorist offence – for example, the suspect has drawn up a plan of the terrorist attack and is actively seeking the means to execute it.

2. USE OF STINGS IN COUNTER-TERRORISM

While the use of 'stings' or 'traps' in policing in the US, the UK, Australia, India and elsewhere is a long-standing practice in, for instance, drug law enforcement and in relation to paedophiles, it is controversial.[1] If a police officer pretends to be a drug dealer and repeatedly offers someone drugs in exchange for money, might not the officer be 'creating crime', supposing the person initially rejects the offer but finally succumbs and hands over cash for the drugs offered? Here, what is meant by 'creation of crime' is not that the manufactured event would not have happened absent the sting (although this is obviously true), but rather that the target did not have the predisposition, the ability and/or the opportunity to commit crimes of that type and, therefore, does not and would not commit them (unless via a sting).

On the other hand, if a police officer pretends to be a child in an online sting operation targeting a suspected paedophile, and the suspect responds in the

manner of a paedophile 'grooming' a child, does so on repeated occasions and then turns up to a face-to-face meeting with the officer believing that they are about to meet a child, then the arrest and subsequent conviction of the suspect is evidently morally justified – especially given the seriousness of the crime of sexual abuse of prepubescent children and the difficulty of acquiring sufficient evidence to convict paedophiles by means of alternative methods. What of the use of stings in counter-terrorism (Sherman 2009)? Let us consider two somewhat different scenarios extracted from two US court cases in order to focus our discussion.

Scenario 1:[2] Employees of a store in 2006 reported to police a video provided to them for conversion to DVD showing six men doing military training and shouting 'Allah Akbar'. A Federal Bureau of Investigation (FBI) informant infiltrated the group and recorded an expressed intention to attack Fort Dix and kill soldiers. The group had a map of Fort Dix and conducted surveillance of it. The FBI informant presented himself to the group as an arms dealer. Members of the group tried to procure weapons from him, such as AK 47s, and were arrested. Five members of the group were convicted of conspiracy to commit terrorist acts; two are serving life sentences, one is serving 33 years.

Scenario 2:[3] James Cromitie is a former drug addict and mental patient in New York with extremist jihadist sympathies. An FBI-paid undercover informant (and convicted fraudster), Shahed Hussain, offered Cromitie $250 000 to fire a rocket-propelled grenade at Stewart Air Base and bomb a New York synagogue. Cromitie initially rejected the offer but agreed after he lost his job. Hussain provided Cromitie with (fake) explosives and drove him to a mosque where he was arrested. He received a prison sentence of 25 years.

Before turning directly to consider these two scenarios, let us set forth some of the key elements of stings. Stings should be distinguished from, firstly, police operations in which a police informant simply observes and provides information in relation to a crime, and, secondly, police operations in which an undercover operative might have minimal participation in a particular crime that has already been initiated and that would have been committed or attempted, irrespective of the actions of the undercover operative. For instance, an undercover operative might provide input into the planning of a crime but do so without initiating the crime, without inducing others to carry it out and without participating in the actual execution of the crime.[4] Rather, stings involve law enforcement providing an opportunity for a suspect to perform a crime in circumstances controlled by law enforcement, and they involve law enforcement offering some form of inducement to a suspect to commit the crime in question. Stings, if well planned and successful, guarantee that there is adequate evidence for a conviction (unlike in the case of crimes that are not under the control of law enforcement). Of course, if stings are to be successful,

then they need to be sufficiently realistic to fool wary, often experienced, offenders.

There are various legally enshrined ethical constraints on stings and accountability measures to ensure that these constraints are complied with (see, for example, Sherman 2009). These constraints vary somewhat from one jurisdiction to another. However, some of these constraints will pertain to the suspect. Although the suspect has not committed that particular offence on that particular occasion since it was a manufactured event, there is an assumption that the suspect would have committed another offence like it on another occasion; therefore, there might be a requirement that the suspect has in fact committed crimes of that type on other occasions. In many jurisdictions, an important accountability mechanism is the requirement that law enforcement be granted a judicial warrant to conduct a sting.

Let us now turn to the analysis of our two scenarios. There are a number of contrasts between Scenario 1 and Scenario 2 in respect of the initial justification for the sting operations, the nature of the inducements, the extent of the participation of the undercover operatives in devising and initiating the crime and, therefore, whether or not the target had a predisposition to commit the crime, and whether or not the undercover operatives were law enforcement officers or persons with a criminal background who were paid to engage in the sting.[5]

Concerning the initial justification for the sting operations, Scenario 1 involved reasonable suspicion of engagement in terrorist activities, for example, the military training video received from a member of the public. By contrast, arguably, Scenario 2 simply involved the exercise of an individual right – specifically, the public expression of an extremist political view. The latter is grounds for further scrutiny of someone but, arguably, not for conducting a sting.

Concerning the nature of the inducement, the undercover operative in Scenario 2 provided a strong financial inducement, for example, $250 000, to an unemployed person. By contrast, the undercover operative in Scenario 1 presented himself as an arms dealer in the context of the members of the group's surveillance of Fort Dix, possession of a map of Fort Dix, and their expressed intention to attack it. In short, the inducement in Scenario 1, but not in Scenario 2, was simply the provision of an element of the means to further the already existing intended goal of the target.

The undercover operative in Scenario 1 inserted himself as a collaborator in a pre-existing terrorist plot and offered to provide a key element of the means to realize it, that is, weapons that might have otherwise been difficult to obtain, given their prior convictions. By contrast, the undercover operative in Scenario 2 provided the target with the plan and know-how – while also providing the target with the means that was not otherwise available to him – that is, the

'fake' explosives – to carry it out. So the role of the undercover operative in Scenario 2 was far greater than in Scenario 1; indeed, arguably, it was necessary in Scenario 2 but not in Scenario 1. Certainly, given the significant role of the undercover operative and the nature of the inducement in Scenario 2, it is questionable whether the target had a predisposition to commit the terrorist act.

In Scenario 1, the undercover operative was a law enforcement officer, whereas in Scenario 2 he was a person with a criminal background who was paid or promised a reduction in his own sentence to engage in the sting. This latter scenario presents potential problems, including the recognition that many criminals are unreliable witnesses; for example, they may fail to record exculpatory conversations and/or commit perjury, and are prone to engage in manipulation or coercion, for example, offer an unacceptable level of inducement to the target. These problems are compounded given the incentives in play, for example, a reduction of his own sentence or a large payment if a conviction is secured. Accordingly, the reliability and, therefore, credibility of the person with a criminal background is less than that of a law enforcement officer and, as a consequence, the testimony provided should have less weight, other things being equal. The problems of reliability and credibility are, of course, significantly mitigated by the fact that the conversations between the undercover operative and the target were recorded, as is typically the case, in both Scenarios 1 and 2.

Given these identified contrasts between the two scenarios, we are evidently entitled to conclude, firstly, that the justification for conducting the sting in Scenario 2 was considerably weaker than for conducting the sting in Scenario 1, and, secondly (and notwithstanding the legal situation[6]), that the target in Scenario 2 (but not the target in Scenario 1) was the victim of a morally unjustified sting. At any rate, having analysed our two scenarios and, as a result, identified some of the salient ethical issues to be analysed in Section 2, we might usefully conclude this introductory section with some discussion of the extent of the use of stings in counter-terrorism operations in the US in particular.

Post-9/11, the FBI (working with local police in a Joint Terrorism Task Force) increased its reliance on stings in its efforts to combat terrorism and, in doing so, conducted a significant number of stings. According to a database of terrorist prosecutions (Norris and Groc-Prokopezyk 2016), during the period from 9/11 to February 2016, there were 580 prosecutions of terrorists, of which 58 per cent were jihadists; 317 of the 580 involved undercover agents or informants. There were 144 Islamic State (ISIS) cases prosecuted in the US during the period from March 2014 to August 2017 (Greenberg 2017). These involved eight attacks, of which two were mass attacks, namely, San Bernardino, California, in 2015 and a nightclub in Orlando, Florida, in 2016. The average sentence handed down was 14.5 years. Of the ISIS cases, 49 per

cent of the suspects were foreign fighters, and 45 per cent of the attacks were on domestic targets. Moreover, 61 per cent of the ISIS cases prosecuted by federal authorities involved undercover agents or informants.

Importantly, according to Norris and Groc-Prokopezyk (2016), there were no successful uses of the defence of entrapment in the cases in their database, that is, that the defendant was tricked or deceived into terrorist action, for example, by unreasonable inducements (including threats) and/or did not have a predisposition to perform the terrorist action. However, they also argue that, in a significant number of these cases, the targets were in fact entrapped according to many of the main criteria for entrapment. This might mean that the legal defence of entrapment was not properly applied or adjudicated, or that the legal defence of entrapment, at least in the US, is inadequate as it stands and, therefore, in need of reform. These legal issues are beyond the scope of this chapter. Instead, let us turn to the underlying ethical issues (Dworkin 1985; Sinnott-Armstrong 1999; Miller and Blackler 2006, Ch. 5; Miller and Gordon 2014, Ch. 11; Miller 2016, Ch. 7).

3. ENTRAPMENT AND ETHICS

While we need to distinguish the conditions definitive of entrapment defences from conditions under which stings might be morally – and ought to be legally – justified, nevertheless, the conditions that *ought to be* definitive of entrapment defences are among the conditions under which stings might be morally – and ought to be legally – justified. So the so-called subjective and objective tests used in the US in relation to the legal defence of entrapment provide a useful initial guide to the discussion of the wider ethical issues raised by stings, including counter-terrorism stings and, in particular, the issues of 'creating crime' and (relatedly) of injustice. Presumably, the target of a successful sting who is convicted of terrorism has been unjustly treated if they did not commit, and would not have committed, an act of terrorism, absent the sting. After all, in these circumstances, the only crime (if crime it is) that has been, or will be, committed is the one manufactured by the sting. Of course, in addition to the problem of the injustice to the target, there is the matter of prevention. The primary purpose of stings is to prevent crime and, in the cases of interest to us here, prevent terrorist attacks. However, if the target of a sting did not and would not have committed an act of terrorism (absent the sting), then obviously the primary purpose of the sting has not been achieved, since no terrorist attack has been prevented (other than, perhaps, the one manufactured by the sting operation).

The subjective test of entrapment focuses on the disposition of the defendant. One problem here is to determine what counts as an adequate evidential basis for the existence of a disposition. If a suspect commits an offence

without any inducement from law enforcement, then the offence is, in and of itself, evidence of a disposition to commit the offence (although this evidence is defeasible – it is possible that there was no disposition, as in some cases of provocation). But if law enforcement provides an inducement to commit the offence, then the existence of a disposition must be demonstrated independently of the inducement, and this can be difficult. However, such evidence of a disposition might be provided, for instance, if the suspect has committed such offences in the past.

There is also a question in play here in relation to the analysis of the concept of a disposition. For instance, if someone has an addiction, then they have a disposition. However, one can have a disposition that one is seeking to control, for example, a paedophile who avoids children.

In the case of suspects who are members of a terrorist organization, perhaps the concept of a *standing intention* should be utilized. The conceptual model here is that of a member of the armed forces in wartime. Combatants, by virtue of their occupancy of a role in an armed force engaged in armed conflict, are reasonably presumed to have a standing intention to kill combatants in the opposing force (and will do so unless the latter intervene to protect themselves by killing the former). Roughly speaking, standing intentions activate immediate intentions which, in turn, cause actions. However, one can have a standing intention without having a relevant immediate intention, for example, a combatant who is eating lunch in a secure building, and one can have an immediate intention without having a standing intention, for example, a husband who intentionally kills his adulterous wife in a fit of anger but without any premeditation. Clearly, terrorist-combatants have, or can be presumed to have, standing intentions by virtue of their role in an armed force. Perhaps members of terrorist organizations operating in well-ordered peacetime settings can likewise be presumed to have a standing intention to kill, or assist in killing, innocent civilians. If so, the difficulty of successfully applying the subjective test would be greatly reduced. It would not, however, deal with the problem of so-called 'lone wolf' terrorists (or others apparently only loosely associated with a terrorist organization), since it is precisely their membership of a terrorist organization (and, therefore, occupancy of a functional role in that organization[7]) that is in question. In relation to our two scenarios, it seems that the targets of the sting in Scenario 1 had a standing intention to commit a terrorist attack whereas, arguably, in Scenario 2 the target did not.

The objective test of entrapment focuses on the inducement; specifically, is the inducement one that it is unreasonable to expect the defendant to refuse? For instance, if law enforcement falsely represents the offence as being legal, then this would be an unreasonable inducement. Again, an 'inducement' might take the form of a threat that an ordinary citizen could not reasonably be expected to resist, for example, a threat to one's life if one does not commit the

offence. On the other hand, offering $250 000 to conduct a terrorist attack (as in Scenario 2) is surely an inducement that an ordinary citizen could reasonably be expected to resist.

While these are clear-cut cases, not all inducements are so easily classified in terms of the reasonable/unreasonable criterion. For instance, what if the 'inducement' is a threat not to one's life but to one's limb? Thus, there is a problem of determining what counts as an unreasonable inducement. Moreover, the objective test (narrowly specified) is not sufficient to defeat entrapment for even if an inducement is one that it is reasonable to expect the defendant to refuse, it might still be the case that the defendant would not have committed this kind of offence absent the inducement. Perhaps, for instance, the defendant would not have had the opportunity to commit the offence. Or, perhaps, to return to the subjective test, the defendant had no prior disposition or, better, standing intention to commit the offence.

The implication of the above discussion of the subjective and objective tests for the defence of entrapment is evidently that a counter-terrorism sting would only be morally justified if the target had a standing intention to commit a serious terrorism offence and if the inducement was one that the target could reasonably be expected to resist. However, as the above discussion has also revealed, these conditions are not sufficient to justify the use of counter-terrorism stings.

Stings can be an effective method in relation to preventing terrorist offences by ensuring would-be terrorists are convicted and receive lengthy sentences for stings can enable terrorists to be convicted who otherwise might not be convicted due to lack of evidence. Perhaps Scenario 1 is an instance of this. On the other hand, stings can simply 'create crime' rather than prevent it. Perhaps Scenario 2 is an instance of this since, arguably, it was unlikely that Cromitie would have committed a terrorist offence absent the sting.

However, even if stings are an effective counter-terrorism method, their use would not be justified if less intrusive or otherwise less harmful methods were available, for example, surveillance. Moreover, the harm in question might consist of a reduction in community trust of law enforcement and, as a result, a reluctance to assist law enforcement in combating terrorism. On the other hand, the use of other methods, such as interception of communications, may well be ineffective in some circumstances due, for instance, to the use of strong encryption by terrorists.

Assuming counter-terrorism stings are effective and less harmful than alternative methods, there are, nevertheless, questions in relation to the appropriate targets of stings and the specific offence types that justify sting operations. Presumably, stings should consist of targeted testing of individuals reasonably suspected of engaging in serious terrorist crimes. Accordingly, if an individual merely exercises their right to free speech by expressing an extremist view,

this is not a sufficient justification to conduct a sting, although it may well be a sufficient justification to monitor the individual.[8] Moreover, a problem arises here in relation to very broad definitions of terrorism and a concomitantly expanding set of terrorism offences,[9] for example, prohibitions on travelling to geographical regions such as Syria. Although an expansion of the set of terrorism offences might be justified in terms of, for instance, reducing the flow of recruits into terrorist organizations, reasonable suspicion that one might engage in such an offence (for example, travel to Syria from Australia) would not justify a sting operation, since the offence is, in itself, not a sufficiently serious one and would not, in and of itself, demonstrate membership in a terrorist organization, let alone participation in a terrorist attack. Further, mere associates of persons reasonably suspected of terrorism should not become the targets of stings, for example, by offering to get them to attend a terrorist training camp. Likewise, stings should not consist of random testing (as opposed to merely monitoring) of members of certain groups, for example, of those attending certain mosques, or of otherwise testing the virtue (that is, without reasonable suspicion of terrorist activity) of ordinary citizens. Individuals have a right to freedom from intrusive state interference, that is, a right not to be subjected to integrity tests if one's actions have not otherwise reasonably raised suspicion of unlawful behaviour. There is a contrast here with those in positions of special trust (for example, police) or those using dangerous equipment (for example, driving cars). Random testing of these groups might well be justified under certain circumstances.

Accordingly, let us assume that counter-terrorism stings ought only to be conducted under the following restricted conditions: they are effective and less harmful than alternative methods, and they target individuals who are reasonably suspected of engaging in serious terrorist crimes. Moreover, it was earlier argued that a counter-terrorism sting would only be morally justified if the target had a standing intention to commit a serious terrorism offence and if the inducement was one that the target could reasonably be expected to resist. Accordingly, the following five conditions suggest themselves as morally justifying the use of counter-terrorism stings (Miller 2016, Ch. 7):

1. The counter-terrorism sting in question is likely to be effective and less harmful than alternative methods, and it targets an individual(s) who is reasonably suspected of engaging in serious terrorist crimes;
2. The suspect has the motive and ability to commit a serious terrorist offence, for example, the suspect espouses extremist violence and has bomb-making know-how;
3. The suspect has a standing intention (and/or an immediate intention) to commit a serious terrorist offence (for example, they occupy a role in a terrorist organization as a bomb-maker, recruiter or trainer[10]), or has

otherwise expressed the standing and/or immediate intention to commit a serious terrorist offence (for example, the suspect has drawn up a plan of the terrorist attack and is actively seeking the means to execute it);
4. The suspect is likely to be presented with, or be able to create (as an individual or by acting jointly with others, for example, other members of a terrorist organization) the kind of opportunity afforded them in the sting scenario, for example, the opportunity to commit a suicide bombing; and
5. The inducement offered is one that the suspect could reasonably be expected to resist.

Note, firstly, that the opportunities to commit a type of terrorist offence might be abundant and the required ability to do so quite minimal, for example, drive a truck into a crowd. Accordingly, in such cases, the burden of the justification for a counter-terrorism sting might lie in (1), (3) and (5) (and (2), in respect of the existence of a motive).

Note, secondly, that, as mentioned above, determining whether or not a suspect had a standing and/or immediate intention (as opposed to a motive) to commit a serious terrorist offence might be difficult in the case of 'lone wolf' terrorists. Or, at least, it is difficult to determine until they commence the initial stage of their planned attack, at which point their immediate intention is manifest, and, therefore, a sting operation, even if possible, would serve no purpose; rather, interception is now not only justified but obligatory. Here, the notion of the initial stage of a planned attack is itself problematic; however, at the very least, we should distinguish the planning stage, including any reconnaissance activities (plotting a terrorist action), the provision of the means to execute the plan, and the execution of the plan (the terrorist action itself). In relation to the last of these, there is a difference between attempting to perform the terrorist action – for example, in the case of a sting, planting what is believed to be a live bomb – from actually performing the terrorist action – for example, detonating a live bomb. Perhaps evidence of planning a terrorist attack is not sufficient to demonstrate the existence of a standing intention (let alone of an immediate intention) to conduct the attack (unless, of course, there is corroborating evidence, such as communicating one's intention to others). However, evidence of planning a terrorist attack and of providing oneself with the means to execute the plan is surely sufficient (other things being equal) to demonstrate the existence of a standing intention to conduct a terrorist attack.

4. CONCLUSION

In this chapter, an ethical analysis of the use of stings or traps in police counter-terrorism operations in liberal democracies and, specifically, in the US has been provided. It has been argued that such operations are morally

justified under certain conditions, notably, that they are effective; there are no less-intrusive methods available; the targets of the sting are known to have the motive, ability and opportunity to commit serious terrorist crimes; and the inducements offered are ones that the targets of the sting could reasonably be expected to resist. An important novel condition that is proposed is that the target in question has a standing intention (and/or an immediate intention) to commit a serious terrorist offence – for example, they occupy a role in a terrorist organization as a bomb-maker, recruiter or trainer – or have otherwise expressed the standing and/or immediate intention to commit a serious terrorist offence, for example, the target has drawn up a plan of the terrorist attack and is actively seeking the means to execute it.

NOTES

1. For discussions of the ethical issues arising from traps or stings, see Dworkin (1985), Sinnott-Armstrong (1999), Miller and Blackler (2006, Ch. 5), and Miller and Gordon (2014, Ch. 11).
2. *United States of America v. Duka*, 671 F.3d 329 (3rd Cir. 2011).
3. *United States of America v. Cromitie* (Williams) – 11-2763.
4. However, it needs to be noted that, in the case of some crimes, such as terrorism in some jurisdictions, planning a crime is itself a crime.
5. There are, of course, other considerations, such as whether the target of the sting merely plotted the terrorist act or actually attempted it. It might be argued that, in Scenario 1, the target merely plotted the terrorist act since they never received the weapons, whereas, in Scenario 2, the target attempted it since he was actually provided with the explosives.
6. In both cases, the convictions of the targets were appealed and the appeals failed.
7. The notion of a functional role in an organization can be considered in terms of the collective end theory of organizational action (see Miller 2010).
8. See Chapters 9 and 10 in this volume.
9. See Chapters 1 and 2 in this volume.
10. Conceivably, an individual might be a member of a terrorist organization and have no intention of committing or contributing to terrorist acts. If so, there would still be a presumption in favour of the individual having a standing intention to committing or contributing to terrorist actions – a presumption that could, nevertheless, be overridden.

REFERENCES

Dworkin, Gerald (1985), 'The serpent beguiled me and I did eat: entrapment and the creation of crime', *Law and Philosophy* 4(1), 17–39.

Greenberg, Karen J. (2017), *Terrorism Prosecutions in the United States: The ISIS Cases March 2014–August 2017*, New York: Center on National Security at Fordham University.

Miller, Seumas (2010), *The Moral Foundations of Social Institutions*, New York: Cambridge University Press.

Miller, Seumas (2016), *Corruption and Anti-Corruption in Policing*, Cham, Switzerland: Springer.
Miller, Seumas and Ian Gordon (2014), *Investigative Ethics: Ethics for Police Detectives and Criminal Investigators*, London: Wiley-Blackwell.
Miller, Seumas and John Blackler (2006), *Ethical Issues in Policing*, Aldershot, UK: Ashgate.
Norris, Jesse J. and Hanna Groc-Prokopezyk (2016), 'Estimating the prevalence of entrapment in post 9/11 terrorism cases', *Journal of Criminal Law and Criminology* 105(3), 649–59.
Sherman, Jon (2009), 'A person otherwise innocent: policing entrapment in preventative undercover counter-terrorism investigations', *Journal of Constitutional Law* 11(5), 1475–1510.
Sinnott-Armstrong, Walter (1999), 'Entrapment in the net?' *Ethics and Information Technology* 1(2), 95–104.

9. Counter-terrorism, social media and the regulation of extremist content

Levi J. West

1. INTRODUCTION

Social media has become a key platform for terrorist organizations to achieve a diverse range of effects, including the distribution of propaganda, the recruitment and radicalization of operatives and sympathizers, and the decentralized facilitation of operations. As social media has come to play an increased role in the terrorist arsenal, it has concomitantly played a more substantial role in counter-terrorism. This role has evolved over time and has seen an increase in prominence for the private sector in a range of functions across the counter-terrorism spectrum.

This chapter articulates some of the various ethical challenges associated with terrorists' use of social media, and the countering of it. It first provides a definition of social media before evidencing some of the mechanisms by which terrorist entities have exploited social media. The chapter then focuses on three key ethical challenges that the increased prevalence of social media in counter-terrorism presents: it looks at the challenge of corporate censorship in relationship to counter-terrorism and the role of social media companies; it discusses the issue of free speech and the impact on political discourse of censoring speech communicated through social media; lastly, it considers the substantial challenge of the different manner in which Salafi-Jihadist content is treated by social media companies, compared to content generated by or sympathetic to the extreme right. In doing so, this chapter emphasizes the inherently multidimensional and contentious nature of contemporary counter-terrorism and the downside costs associated with its prosecution.

2. TERRORISM AND SOCIAL MEDIA

Terrorism has always retained a communicative dynamic and, by extension, a necessarily dependent relationship with information and communications technology (ICT). Since the emergence of modern terrorism in Tsarist Russia

in the late 1800s, terrorism has been reliant on ICT to perpetuate itself and to propagate the ideological propaganda either associated with or embedded in its violence. The relationship between terrorism and ICT has evolved along with the character of ICT, while the nature of the relationship has endured. Russian and European anarchists exploited the printing press and the emergence of the industrialized news media (Laqueur 1977), Palestinian terrorist organizations exploited the emergence of international satellite television broadcasting capacity, and the Islamic State (ISIS) exploited the emergence and increasing ubiquity of social media platforms (Awan 2017; Farwell 2014). The use of social media by terrorist entities should be understood in the context of a broader appreciation of the centrality of communicative considerations to terrorist violence, and the evolution of the relationship between terrorism and ICT.

Two key incidents demonstrate the significance of social media to contemporary terrorist operations and the manner in which the counter-terrorism apparatus has become obligated to evolve its approach. The first is the release, in August 2014, of the beheading video of American journalist James Foley. The video was highly choreographed, with the victim dressed in an orange jumpsuit in reference to those detained at the Guantanamo Bay Naval Base and to the previous beheading of Nicholas Berg by Abu Musab al-Zarqawi in 2004. In addition, the executioner spoke with a thick British accent. Most importantly, the video was initially distributed on the social media site Diaspora before being uploaded to LiveLeak, and ultimately, and repeatedly, to YouTube, with a Twitter campaign to encourage sympathizers to upload the video and share it with fellow supporters of ISIS (Friis 2015). The release of this video resulted in then President Obama holding a press conference in which he condemned ISIS, and it ultimately resulted in the establishment of the 30-country coalition Combined Joint Task Force – Operation Inherent Resolve, whose mission was to:

> defeat ISIS as a military force on the battlefield in Iraq and Syria. We will disrupt their ability to command and control their fighters, remove their safe havens, interrupt their revenue streams that fund their operations, destroy their equipment, and kill their fighters. We will eliminate their effectiveness as an organized force on the battlefield. (Combined Joint Task Force Operation Inherent Resolve 2014)

The second incident that highlights the contemporary power of social media in the terrorism and counter-terrorism ecosystem is the Christchurch terrorist attack in March 2019. In that incident, the Australian perpetrator livestreamed his mass shooting at two mosques via Facebook (Macklin 2019). In addition to the livestream, the shooter shared a manifesto, *The Great Replacement*, via Twitter and 8chan. Facebook announced that they removed 1.5 million

instances of the video from their platform in the 24 hours after the attacks (Mahtani 2019). Subsequently, there have been at least three terrorist incidents by extreme right-wing actors that have explicitly made reference to having been inspired by the Christchurch attacker – two in the United States and one in Norway. The video itself, the ideas contained in the manifesto, and the actions of the Christchurch shooter have all become animating characteristics of extreme right-wing online discussion forums across numerous platforms, as have the actions of those he inspired.

These two incidents demonstrate the powerful platform that social media provides for terrorist entities, movements, and actors. The global reach, near ubiquity, and non-existent barriers to entry allow social media to be both a powerful offensive tool for terrorists and a powerful means of creating, sharing, and accessing extremist content for supporters and sympathizers. Social media technology is also ideologically agnostic, as the above incidents demonstrate, and is equally as powerful to those motivated by Salafi-Jihadist ideas as it is to those motivated by extreme right-wing ideas. It is also important to recognize that the exploitation of social media by terrorist entities has not come about by accident. Both the jihadist movement (West 2016) and the extreme right (Conway et al. 2019b) have embraced strategies that are ideally suited to the utilization of online communication technologies, and, as identified above, the use of emergent ICT is now at the core of any terrorist campaign. While social media is the modern manifestation of a problem that has persisted since at least the days of anarchist terrorism in the late 1800s, it has produced a range of counter-terrorism innovations as nation states have sought to counter its use.

Before articulating the novel ethical challenges raised by the counter-terrorism practices that have emerged in response to the use of social media by terrorist entities, it is necessary to provide an appropriate set of definitional parameters. For the purposes of this analysis, social media will be defined as it is in the *Dictionary of Social Media*, which conceptualizes it as: 'The online and mobile technologies or platforms people use to interact and share content, including social networking sites, social bookmarking and social news sites, geosocial networking sites, blogs, online forums, file-sharing and media-sharing sites, social gaming sites, social commerce sites, virtual worlds, and wikis' (Chandler and Munday 2016).

This definition of social media has the benefit of being sufficiently broad as to incorporate the initial social media platforms, such as YouTube, Facebook, and Twitter, while also providing sufficient scope to consider more recent manifestations, such as 4chan, 8chan, and Discord. One of the core challenges that terrorist and extremist use of social media has presented has been the relative ease with which these actors have been able to shift platforms in response to countermeasures and de-platforming efforts. The dynamic nature of social

media itself, as well as the relatively short lifespans of current engagement by terrorists and extremists with specific platforms, means a broad conceptualization is necessary.

Overwhelmingly, irrespective of ideological persuasion or platform preference, terrorists and extremists utilize social media to achieve three core objectives. As I argued in '#jihad: Understanding Social Media as a Weapon', ISIS's use of social media serves to 'reinforce their narratives to multiple audiences, contributes to recruitment and radicalisation, and of most consequence to Western security agencies, is increasingly responsible for substantially contributing to terrorist attacks in Western countries' (West 2016, pp. 9–10).

It is necessary to recognize that, while the distribution of propaganda and the concomitant process of recruitment and radicalization of individuals and small cells has been a function of terrorist and extremist exploitation of ICT since the 1800s, the refined, targeted, and accessible decentralized command and control campaign of ISIS from 2014 onwards was unprecedented in its scale, sophistication, and effectiveness. Bergema and Kearney (2020) identified 116 terrorist attacks against Western jurisdictions between 2004 and the end of 2019 and note that '[t]errorist violence spiked between mid-2014 and late 2017' (p. 10). It was during this period that ISIS's external operations capability was at its most effective. A number of scholars have articulated the mechanisms by which this system operated (West 2016; Meleagrou-Hitchens and Hughes 2017; Nesser et al. 2016), and within policy environments and public discourse, ISIS became synonymous with lone-actor terrorist attacks undertaken by individuals who had accessed online material via social media. It was in response to this significant increase of terrorist activity (noting that the date cited here excludes those operations that were disrupted by Intelligence services without being detailed publicly) – that witnessed 25 terrorist attacks in the West in 2016, with approximately 27 in 2017 – that counter-terrorism operations and policies undertook a substantial recalibration in regard to social media and its effective exploitation by ISIS.

The sudden and dramatic expansion and increased effectiveness of ISIS's social media efforts resulted in a range of novel counter-terrorism responses that continue to evolve in parallel with the threat. In September 2014, ISIS released a now infamous audio statement from their then spokesperson and head of external operations, Abu Mohammed al-Adnani titled 'Indeed Your Lord is Ever Watchful'. The statement was released in Arabic and English, and followed up with a transcript of the speech in the fourth edition of their English-language digital magazine *Dabiq*. At the time, these pieces of operationally oriented propaganda were released with relative impunity on Twitter and other social media platforms. In contrast, in 2019, Brian Fishman, a terrorism scholar and now head of counter-terrorism policy at Facebook, wrote that 'Social media companies, both individually and in concert with one another,

have developed robust operations to prevent terrorists from abusing their platforms. Like all counter-terrorism programs, these efforts are imperfect, but they represent a significant new component of the societal response to terrorist violence' (Fishman 2019).

3. SOCIAL MEDIA INTELLIGENCE (SOCMINT)

Before the escalation of usage of social media by terrorist entities manifested as severely as it did from 2014 through 2017, there were already policy-oriented discussions of the utility of social media for Intelligence purposes. In 2012, Sir David Omand, the former director of the Government Communications Headquarters (GCHQ), the British Signals Intelligence service, co-authored a landmark paper that advocated for the exploitation of social media and social networks for Intelligence purposes. The paper not only provides an articulation of the opportunities that increased the use of social media presence, but also highlights some of the risks associated with these developments. In articulating the case for exploiting social media for Intelligence purposes, the authors state:

> social media intelligence – which we term 'SOCMINT' – could contribute decisively to public safety: identifying criminal activity; giving early warning of disorder and threats to the public; or building situational awareness in rapidly changing situations. As society develops and adopts new ways to communicate and organise, it is vital that public bodies, including law enforcement and the intelligence community, keep up with these changes. (Omand et al. 2012, p. 9)

In the intervening years, the exploitation of social media for Intelligence purposes became increasingly interwoven into the fabric of counter-terrorism Intelligence practices, but it was the campaign by ISIS that drove the first major incorporation of active, offensive private-sector activity in the social media domain. The initial formalization of this shift is reflected in the establishment of the Global Internet Forum for Countering Terrorism (GIFCT), established by Facebook, Microsoft, Twitter, and YouTube in July 2017. This forum provides an ongoing platform for the sharing of knowledge and capability regarding counter-terrorism on social media platforms by private-sector technology companies. It is reflective of the shift in both the acknowledgement of the problem and the prioritization of countermeasures that four of the largest technology companies in the world felt it necessary to establish an entity such as the GIFCT. It is also reflective of the increasingly central role of private-sector actors in the management and mitigation of access to what is deemed extremist material in the online environment. It is this evolution of counter-terrorism policy and practice that presents a number of ethical challenges and is the focus of the balance of this analysis.

4. THE ETHICAL ISSUES

The emergence and institutionalization of private-sector activity specific to counter-terrorism activity raises a number of ethical issues that are worthy of consideration. I note that, while these practices raise substantive ethical issues, the discussion herein is not intended as a blanket condemnation of these practices, nor is it a suggestion that they ought to cease. The measures that have been developed reflect the evolving and dynamic nature of the contemporary terrorist threat, and active engagement in countermeasures by companies such as Facebook and Twitter are a necessary component of contemporary counter-terrorism. The practices do, however, raise substantial ethical questions about the role of the state, the role and purpose of technology companies, and the place of free speech within liberal democracies as it pertains to extremist content and ideology.[1] Additionally, as extremist and terrorist content generated by entities and individuals sympathetic to extreme right-wing ideology becomes increasingly prevalent, it is important to highlight the distinct and more challenging problem presented by this further evolution in the nature of terrorist exploitation of social media.

5. PRIVATE-SECTOR CENSORSHIP

The first major ethical issue that arises from private-sector counter-terrorism activity pertains to the role and purpose of private-sector, for-profit companies as distinct to the role of the state. The increased role of private companies has manifested in a range of policies and practices, but central to this has been what has become known as takedowns (Conway et al. 2019a). This involves the identification of content deemed extremist or terrorist related, and that meets the various criteria established by the companies for the content to be removed from their platforms. Additionally, user accounts can be suspended or permanently blocked. Distinct from the assumptions implicit in the approach and any assessment of the effectiveness of this activity, there are substantial differences between the forms of accountability that private companies have when considered against the state,[2] in particular with regard to its role as censor.

A private company has, as one of its main institutional purposes, the generation of profit. Other obvious central purposes include the provision of some good or service to the community and of employment to its managers and workers. The shareholder value theory holds not only that its overriding institutional purpose is and ought to be to maximize shareholder wealth (via maximizing profit), but also that it is a legal requirement that corporations, in particular, have this as their overriding purpose. However, this latter claim is

false, and its underpinning normative 'theory' is highly controversial to say the least (Stout 2012).[3] Moreover, we need to distinguish between the de facto goals of an individual company and the normative goals of the market-based industry in which it competes. According to the standard normative theory of markets, the goal of an industry or market is to provide a social benefit *in part* by means of the so-called invisible hand, therefore, regulatory arrangements need to be put in place to ensure that the social benefits in question are, in fact, forthcoming. At any rate, the institutional purposes, whether de facto or normative, of companies are quite different to those of the state (Miller 2010). Specifically, the state has as one of its primary purposes to ensure the safety and security of members of the public. By contrast, the engagement of private-sector companies such as Facebook, Twitter, and the like in counter-terrorism activity is not undertaken, generally speaking, because their purpose is to protect the public from extremist or terrorist content, but, at least in large part, because there is concern that the ongoing presence of that content on their platforms potentially diminishes their profits, or that their profit-generating activities will be negatively impacted by coercive action by the state.

Additionally, private-sector entities, such as Facebook or Twitter, lack even rudimentary accountability in comparison with liberal democratic states. Unlike a government that, should they be deemed to have failed to protect their populace from extremist or terrorist activity, may be removed from office through electoral processes or suffer similar political repercussions, private companies that fail to take appropriate action against extremist or terrorist content have relative impunity.[4] Current debates regarding Facebook and its responsibility to police extreme right-wing content on its platform is reflective of this dilemma (Knaus et al. 2019; Timberg 2020; Love 2019). In the strictest of senses, Facebook has no obligation to police this content, at least by reference to it being deemed offensive to the broader public.

By contrast, the liberal democratic state does have obligations with respect to policing content. Indeed, it has the role of censor – which it typically exercises through independent authorities, such as censorship boards and the like – however, it does so only: (a) via its legislature, which ought to reflect the standards/views of the citizenry, and (b) via the constitution or otherwise underlying structure of individual rights (for example, right to freedom of speech and freedom of the press). In short, while the government ought to do the censoring (or, at least, its established independent boards ought to do so), the principles it applies ought to reflect community standards (democracy) and individual rights (liberalism).

However, the cost of an expansion of private-sector policing of extremist content has been the withdrawal of the state from its role in expressing and, at times, upholding community-derived norms and standards with regard to what

constitutes unacceptable material. I note here that liberal democratic states mark themselves out as different from authoritarian states because liberal democratic states do not determine the good life for the people, are constrained by democratic principles, and have a commitment to individual rights.

Recognizing these limits and processes in liberal democracies, the state remains highly active in the regulation of pornographic materials, child-exploitation material, and other forms of offensive content. The state has become increasingly dependent, in part for technical and scale reasons, on the private sector to police itself in relation to extremist content. This raises a number of ethical challenges related to whether for-profit corporations are the best mechanism for society to determine what content constitutes extremist material, in what levels of tolerance of dissenting and uncomfortable ideas society is prepared to engage, and how the production and distribution of extreme or dangerous ideas by powerful and influential actors should be treated and managed within and across society. The absence of effective accountability in relation to private companies results in circumstances whereby, should society deem the actions of a private company to be out of step with the broader norms of the populace, the state lacks meaningful mechanisms of redress.

6. CENSORING EXTREME IDEAS?

An extension of the changing role of the state and the increased role of private companies in counter-terrorism, in particular with relation to extremism, is the challenges that the policing of extremist content, as distinct from terrorist content, has with regard to freedom of expression. While the state may remain committed, at least in principle, and in certain jurisdictions in law, to the principles of freedom of expression, increasingly, private companies exercise authority over core public domains and platforms of public discourse. As a result, the norms surrounding freedom of expression, as determined by an unaccountable private company have a substantial role in the policing of content and material that may be deemed unacceptable according to the terms and conditions of the company, rather than the representative determinations of a democratically elected government and its public service.

This has further implications and raises further ethical challenges with regard to the aforementioned institutional purpose of a private company as compared with the state, and it has substantial implications for increasingly diverse sets of ideological motivations that produce and share extremist content. The basis on which a private company makes assessments as to what constitutes extremist content is categorically different to that of the state, and is anchored in breaches of terms-of-service agreements that, in turn, reflect the

norms of the company as opposed, necessarily, to the norms determined by democratic processes.

7. CENSORING POLITICAL DISCOURSE

The resurgence of the extreme right has presented a unique challenge for Western democratic states. Unlike Salafi-Jihadism, which is not an extension of Western political discourse, extreme right-wing ideas related to immigration, demographics, globalization, and race are all extensions of mainstream political discourse within Western democratic politics. This results in a much more complex exercise in identifying those aspects of contemporary political discourse that can be classified as extreme, and those that merely exist at the edges of what can be considered acceptable political discourse.[5] This more challenging dynamic exists regardless of whether it is the state or private companies seeking to police said discourse. The challenge is greatly exacerbated by the increased institutional representation and propagation of extreme ideas, and their broadcast, reinforcement, and propagation by powerful institutional political actors, including presidents, prime ministers, senior elected officials, and through policy determinations. As an example, in the aftermath of the Christchurch attack – committed by an Australian citizen in 2019 – Australian senator Fraser Anning tweeted, 'Does anyone still dispute the link between Muslim immigration and violence?'[6] followed by a statement that stated, in part, 'the real cause of bloodshed on New Zealand streets today is the immigration program which allowed Muslim fanatics to migrate to New Zealand in the first place' (Press Association 2019).

The comments on Twitter remain accessible and, as of the time of writing, former Senator Anning's account remains active. While Senator Anning did not explicitly endorse or call for violence, his comments were sufficient to receive formal censure from the Australian Senate. While it is not necessarily Twitter's obligation to police Australian political discourse, or any other jurisdiction's political discourse, these comments, and the litany of parallel examples within political discourse across democratic states, serves to highlight the shortcomings of private companies being obliged to police certain manifestations of political discourse while other manifestations remain accessible.[7]

While platforms such as Facebook and Twitter have taken substantial steps towards actively addressing the presence of Salafi-Jihadist content on their websites, the challenge of the extreme right remains a much more complex environment. In 2020, Facebook took the bold step of removing ads purchased by the Trump re-election campaign after it was revealed they contained Nazi symbols (Karni 2020). The differential response to extreme right-wing content is as much a function of the difficulty of discerning explicitly extremist content from far-right content as it is a reflection of the views of right-wing extremists

increasingly being reflected by political actors in positions of power. The concern here is that powerful corporate entities may potentially be affected by government decision-making should they act in ways deemed to be against the interests of politicians with an interest in extreme right-wing political discourse or advancing their own political agenda.

8. OFFLINE PROPAGATION

Beyond these substantial ethical challenges, the policing of extremist and terrorist content suffers from a number of broader issues. For instance, at the core of the activity is an assumption that the removal of content is an effective countermeasure. This is undermined by a number of key considerations. In the first instance, the proactive measures undertaken by the major technology companies have resulted in the majority of extremist and terrorist activity having shifted to encrypted platforms, or platforms that they perceive to be encrypted. As a result, the content in question remains accessible, albeit not as easily accessible. Additionally, extremist and terrorist content is not a necessary requirement for individuals to be exposed to extremist ideas or to undertake a terrorist attack. As any consideration of the long history of terrorism demonstrates, terrorism was a security challenge well before the advent of the Internet or social media, and eradicating extremist and terrorist content will not eradicate terrorism. The point here is not that private and state institutions should do nothing to limit accessibility or to make it more difficult for terrorist organizations to propagandize, recruit and radicalize, and facilitate operations. Rather, it is that social media is not a necessity for terrorist activities to be undertaken. The historical connections between terrorism and propaganda bear this point out.

9. CONCLUSION

As highlighted above, the development and refinement of counter-terrorism policies and practices by social media and technology companies has come about as a function of the evolving nature of terrorist and extremist entities' exploitation of said platforms. The challenges highlighted above, namely, the distinct institutional purposes, the impacts of private companies acting as determinants of freedom-of-expression norms, and the disparate treatment of Salafi-Jihadist content compared with extreme right-wing content are a sample of the kinds of ethical dilemmas that are raised by the changing character of counter-terrorism activity.

As these practices continue to evolve, it is important that the state continues to retain its role in expressing and representing community norms and as an accountable actor in the policing of freedom of expression. Ensuring that an

increasingly diverse body of extremist content is responded to with equal vigour, and provided with equally limited tolerance, is a key role for the state. The increased presence of private-sector technology companies as controllers and owners of key public and political discourse infrastructure presents a number of substantial risks to the traditional roles that the democratically accountable apparatus of the state played in ensuring civil political discourse was both enabled and protected. As technology plays an increased role in politics and in political discourse, it is essential that the ethical challenges that are raised by these changes are considered and incorporated into policy thinking, both within government and within private-sector technology companies.

NOTES

1. For more on this, see Henschke's Chapter 10 in this volume.
2. In referring to 'the state' here, it is a reference to variations of liberal democratic states, with some form of relatively effective electoral accountability.
3. For more on the ethico-normative discussion of institutions, see Chapter 10 of Miller (2010).
4. However, we must recognize that private institutional behaviours can be changed through shareholder activism and direction, consumer boycotts, social criticism, criticism by politicians/political leaders, and so on. There are various mechanisms outside of specific legislative changes that can alter private institutional behaviours and practices.
5. As noted by Henschke and Reed, identifying what counts as extremism is contested and raises significant ethical complexities (see Henschke and Reed forthcoming).
6. Fraser Anning, Twitter post, 15 March 2019, https://twitter.com/fraser_anning/status/1106432499704451072.
7. See Henschke and Reed (forthcoming) for more on the ethical challenges around this.

REFERENCES

Awan, Imran (2017), 'Cyber-Extremism: Isis and the Power of Social Media', *Society* 54 (2), pp. 138–49. doi: 10.1007/s12115-017-0114-0.

Bergema, Reinier and Olivia Kearney (2020), 'Rise O Muwahhid, Wherever You May Be: An Analysis of the Democratization of the Terrorist Threat in the West', *ICCT Report*, The Hague: International Centre for Counter-Terrorism.

Chandler, Daniel and Rod Munday (2016), 'Social Media', in Daniel Chandler and Rod Munday (eds), *A Dictionary of Social Media*, Oxford: Oxford University Press.

Combined Joint Task Force Operation Inherent Resolve (2014), *Combined Joint Task Force Operation Inherent Resolve Fact Sheet*, https://www.inherentresolve.mil/Portals/14/Documents/Mission/History.pdf?ver=2016-03-23-065243-743.

Conway, Maura, Moign Khawaja, Suraj Lakhani, Jeremy Reffin, Andrew Robertson and David Weir (2019a), 'Disrupting Daesh: Measuring Takedown of Online Terrorist Material and Its Impacts', *Studies in Conflict & Terrorism* 42 (1–2), pp. 141–60.

Conway, Maura, Ryan Scrivens and Logan Macnair (2019b), 'Right-Wing Extremists' Persistent Online Presence: History and Contemporary Trends', *ICCT Policy Brief*. doi: 10.19165/2019.3.12.

Farwell, James P. (2014), 'The Media Strategy of ISIS', *Survival* 56 (6), pp. 49–55. doi: 10.1080/00396338.2014.985436.

Fishman, Brian (2019), 'Crossroads: Counter-Terrorism and the Internet', *Texas National Security Review* 2 (2), pp. 82–100. doi: 10.26153/tsw/1942.

Friis, Simone Molin (2015), '"Beyond Anything We Have Ever Seen"': Beheading Videos and the Visibility of Violence in the War against ISIS', *International Affairs* 91 (4), pp. 725–46.

Henschke, Adam and Alistair Reed (forthcoming), 'Ethics and Responses to Extremist Propaganda Online: Untested Assumptions, Key Responses, and Ongoing Considerations', *Studies in Conflict and Terrorism*.

Karni, Annie (2020), 'Facebook Removes Trump Ads Displaying Symbol Used by Nazis', *New York Times*, 18 June, https://www.nytimes.com/2020/06/18/us/politics/facebook-trump-ads-antifa-red-triangle.html.

Knaus, Christopher, Michael McGowan, Nick Evershed and Oliver Holmes (2019), 'Inside the Hate Factory: How Facebook Fuels Far-Right Profit', *The Guardian*, 6 December, https://www.theguardian.com/australia-news/2019/dec/06/inside-the-hate-factory-how-facebook-fuels-far-right-profit.

Laqueur, Walter (1977), *A History of Terrorism*, New Jersey: Transaction Publishers.

Love, David A. (2019), 'It Is time for Facebook to Stop Coddling the Far Right', *Al Jazeera*, 21 December, https://www.aljazeera.com/indepth/opinion/time-facebook-stop-coddling-191216131255890.html.

Macklin, Graham (2019), 'The Christchurch Attacks: Livestream Terror in the Viral Video Age', *CTC Sentinel* 12 (6). https://ctc.usma.edu/christchurch-attacks-livestream-terror-viral-video-age/.

Mahtani, Shibani (2019), 'Facebook Removed 1.5 Million Videos of the Christchurch Attacks Within 24 Hours—and There Were Still Many More', *Washington Post*, 17 March, https://www.washingtonpost.com/world/facebook-removed-15-million-videos-of-the-christchurch-attacks-within-24-hours--and-there-were-still-many-more/2019/03/17/fe3124b2-4898-11e9-b871-978e5c757325_story.html.

Meleagrou-Hitchens, Alexander and Seamus Hughes (2017), 'The Threat to the United States from the Islamic State's Virtual Entrepreneurs', *CTC Sentinel* 10 (3), https://ctc.usma.edu/the-threat-to-the-united-states-from-the-islamic-states-virtual-entrepreneurs/.

Miller, Seumas (2010), *The Moral Foundations of Social Institutions: A Philosophical Study*, Cambridge: Cambridge University Press.

Nesser, Petter, Anne Stenersen and Emilie Oftedal (2016), 'Jihadi Terrorism in Europe: The IS-Effect', *Perspectives on Terrorism* 10 (6), pp. 3–24. https://www.jstor.org/stable/26297702.

Omand, David, Jamie Bartlett and Carl Miller (2012), *#Intelligence*, New York City: Demos.

Press Association, 'Fury as Australian Senator Blames Christchurch Attack on Muslim Immigration', *The Guardian*, 16 March, https://www.theguardian.com/world/2019/mar/15/australian-senator-fraser-anning-criticised-blaming-new-zealand-attack-on-muslim-immigration.

Stout, Lynn (2012), *The Shareholder Value Myth: How Putting Shareholders First Harms Investors, Corporations and the Public*, Oakland: Berrett-Koehler.

Timberg, Craig (2020), 'How Conservatives Learned to Wield Power inside Facebook', *Washington Post*, 21 February, https://www.washingtonpost.com/technology/2020/02/20/facebook-republican-shift/.

West, Levi J. (2016), '#jihad: Understanding Social Media as a Weapon', *Security Challenges* 12 (2), pp. 9–26.

10. On free public communication and terrorism online

Adam Henschke

1. INTRODUCTION

The basic issue that this chapter examines is how we set morally justifiable limits on what people say online, with a specific focus on speech acts related to terrorism. The chapter takes as one of its foundational assumptions that there are situations in which free speech, while an important and potentially fundamental right, can be justifiably limited: 'If there is a Free Speech Principle, it means that free speech is a good card to hold. It does not mean that free speech is the ace of trumps' (Schauer 1982, p. 9). This chapter draws from discussions of free speech to explain how and when particular sorts of constraints on public communications are justified. It begins with a recognition that the norms of behaviour and free speech online are evolving, moves to an exploration of the key conceptual and ethical discussions around free speech, and closes with a set of framing questions to give some guidance as to how we can approach the question of whether online terrorist public communications should be restricted or not.

2. TERRORIST USE OF THE INTERNET

One significant recent counter-terrorism challenge is concerned with how terrorists use cyberspace as a tool for their activities, particularly social media, to distribute their messages and to recruit individuals to their causes. One response is to constrain online activity, but this runs up against the idea of free speech, and the centrality of free speech to online behaviour. In its earlier days, the anonymity of the Internet afforded near-absolute freedom of speech (Qasir 2013, p. 3667) and was hoped to liberate us and spread democracy though the world. The Internet 'flattens the world', and this 'helps because it frees up people and capital to different, more sophisticated work' (Friedman 2005, p. 21). Social media was expected to be a further aid for individual freedom. The capacity to publicly criticize and shame people 'was coercive, borderless

and increasing in speed and influence. Hierarchies were being levelled out. The silenced were getting a voice. It was like a democratization of justice' (Ronson 2015, p. 9). Such thoughts about the Internet and free speech reached their zenith with the Arab Spring, where social media were thought to be a key part of revolutions (Shehabat 2012). These uprisings largely fell flat, and the idea of social media being uncontested democratic goods approached its nadir (Morozov 2013). The so-called Islamic State (ISIS) were soon using social media to broadcast beheadings and their military takeover of large portions of Iraq and Syria (Kilcullen 2016; Klausen 2015). They made direct threats to people's safety in areas geographically distant from the main theatre of conflict (ABC News 2016) and used social media to recruit people to their cause (Gates and Podder 2015).

The willingness and apparent skill at which terrorists like so-called ISIS used social media as a tool for communication puts pressure on how *free* speech ought to be. At the same time, we see individuals asserting that Twitter and Google actively censure the so-called alt-right movement in the US (Leetaru 2018), and in mid-2020, Twitter started adding 'fact-checking' notes to the US president's tweets. Cyberspace is a complex environment, and navigating free speech debates as they evolve and apply to it is complex and challenging.

This chapter focuses on free speech and its relevance to terrorist use of the Internet – specifically, terrorist use of the Internet to promote their cause and recruit people to their cause. In keeping a focus on the idea of free speech, this chapter will not look at terrorist use of the Internet for planning and running operations, nor encrypted and constrained communications between selected individuals, but on speech acts intended for 'public consumption'. The intention is to expose some of the relevant 'moral mechanics' of free speech and terrorist use of the Internet, and to help explain why certain moral norms can and should be promoted for cyberspace.

2.1 The Idea of Free Speech and Liberal Democracies

While the modern notion of free speech has its roots in European thought,[1] much of its modern form and recent evolution draws from debates in the US (Sadurski 1999, p. 5). The First Amendment states that 'Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances'. Certain speech acts pertaining to terrorism may be protected by claims of free speech, while other speech acts pertaining to terrorism might be justifiably constrained.

In the US and elsewhere, the notion of free speech is often considered a central fact of liberal democracies,[2] and following John Stuart Mill, is

thought of as constitutive of liberal democracy. 'The right to free speech is hardly in tension with democracy; it is a precondition for it' (Sunstein 1993, p. 121). While different nations and people conceptualize free speech differently, and draw the limits differently,[3] liberal democracies mark themselves as liberal in part by the fact that they take free speech seriously (Sadurski 1999, p. 1). This may have different foundations. Free speech may be necessary for citizens to develop, grow, and become themselves (Sadurski 1999, p. 17), or be fundamental to self-development; it is a basic right held by individuals in virtue of their need to flourish (Griffin 1986, p. 229). For others, it is simply a function of liberal democracies; the state has no place in telling people what they believe and can say. To constrain free speech is to violate self-expression, and state curbs on self-expression are inimical to the liberal democratic ideal that its citizens can pursue whatever notion of the good life they desire (Larmore 1987, p. x).

2.2 Limits in Liberal Democracies and Terrorism

However, liberal democracies recognize that free speech can be harmful. 'Speech is plainly not a self-regarding act…Affecting others is most often the whole point of speaking…[S]peech clearly can and frequently does cause harm' (Schauer 1982, p. 10). This claim should be clear when thinking of terrorism – core to any conception of terrorism is the notion of targeting innocents for harm in order to achieve some political, social, or ideological end.[4] The terrorist action can itself be considered a communicative act, a performance of violence to further the political, social, or ideological end.[5] This communicative aspect of terrorism should be considered central to any successful terrorist activity. 'The success of a terrorist organisation depends almost entirely on the amount of publicity it receives…it is not the magnitude of the terrorist operation that counts but the publicity' (Laqueur 1977, p. 109). The recent evolution of international terrorism has seen their active and skilful use of social media as part of the communicative aspects of terrorism (Awan 2017).

To explain the moral mechanics of free speech, we can start with relatively simple examples where constraints on free speech seem justified, to draw out particular notions, concepts, or values that are frequently seen to play a role in justified constraints on free speech. Constraints on child pornography[6] are a useful device to start seeing the limits of an argument around free speech. We might, for example, be opposed to the production, distribution, and/or viewing of child pornography because such actions do not show respect for the victims. This basic notion, that certain speech acts can be constrained, can be explained by reference to the notion of offence (van Mill 2017).[7] '[W]e can say that what is wrong with giving offence in general is that it is showing a lack of respect for

others and that it may cause them to lose some of their self-respect' (Weckert 2007, p. 29).

An analogy can be made between child pornography and certain terrorist communications. In addition to the obvious significant moral wrong of beheading someone, videos where members of so-called ISIS behead people ought to be constrained because they show a lack of respect for the victims of terrorism. This includes the person who was beheaded, their loved ones, and, potentially, other victims of IS's crimes and so on. Given the obvious significant moral wrong of the violent act itself – like child pornography – and the offence that public communication of the violent act causes, such communications ought to be constrained. A further point is that what gives justification for constraints on child pornography and terrorism is not simply 'mere' offence. Rather, the offence has to be *significant* and *explicable* (van Mill 2017).

Another situation where a speech act could potentially be legitimately constrained is where that speech act is likely to result in significant *harms*. The idea here is that speech acts can cause actual physical harm or induce people to violence, and so the speech act ought to be constrained. As US Justice Oliver Wendell Holmes stated in *Schenck v. United States* (1919), 'The most stringent protection of free speech would not protect a man falsely shouting fire in a theater and causing a panic'. We see that the same principle applies with certain terrorist communications – production, distribution, and owning terrorist material/manuals and so on may be justifiably constrained if use of that material enables one to cause physical harms. As with offence, a fundamental point is that what gives justification for constraints on incitement and terrorism is not simply 'mere' harm – the harm has to be *significant* and *likely* (van Mill 2017).

A more complicated scenario involves 'hate speech'. This is a more contested space than significant offence and harm. In some ways, hate speech is a hybrid between offence and harm, however, the offence and harm are relative to a particularly vulnerable group. If a particular group has already or historically suffered significant harm, and the given speech act either recalls this or does not show proper respect for this, this group's capacity to be offended is greater. A particular group has to be singled out/targeted – either explicitly or implicitly – in a particular manner for some act of violence, and so on (van Mill 2017). Hate speech recognizes that history and context matter – offence and harm do not tell the whole story.

I suggest that the relevant functional feature that causes complication and argument around hate speech is an issue of fairness or lack of equivalence: 'Why do I get criticized and threatened with punishment for calling that group an offensive term, if the same group doesn't get criticized or punished when they use it? It isn't fair!' One idea underpinning hate speech is that people deserve different treatment (Fish 1994). It is a special or controversial notion

because it is reliant on one set of people's rights being differently constrained in light of another set of people's rights. This different treatment goes against the notion of fairness as equivalence. 'The aim [of hate speech laws] is simply to diminish the presence of visible hatred in society and thus benefit members of vulnerable minorities by protecting the public commitment to their equal standing in society against public denigration' (Waldron 2009, p. 1600). When considered distinctly from offence and harm, hate speech turns, in part at least, on the notion of special vulnerability and the role that historical treatment of particular groups by other groups plays in that vulnerability.

The relevance of hate speech to terrorism is when there is a context of heightened security fears. If I was to say, 'We ought to stab those police', most people would likely pay no attention to me making such claims. But if this was said minutes after a terrorist attack on police involving knives in the same street, then the heightened threat might justify constraints on such statements at that time. Like hate speech, when thinking of constraints on certain speech, the context matters: speech acts may need to be considered differently depending on the context of the speech act, and active acts of, or incitement to, terrorism involve a relevant context.

The simple principles are that, in situations of significant offence, high chances of significant harm, and/or in a context of heightened threat, we might justifiably place constraints on terrorist speech acts. These principles provide the basis for a deeper moral discussion of constraints around hard cases. We begin with the idea that certain speech acts are justifiably prohibited – producing and distributing beheading videos, instructional material on how to make bombs, incitement to violence against police at times of heightened threat, and so on. There is agreement that these acts ought to have constraints on their *public communication*.[8] However, when looking at terrorism online, there is a large set of material and speech acts that do not easily fit in the offence, harm, or context aspects. For example, ought a person be banned from Facebook for 'liking' a beheading video, tweeting that beheading videos are good, posting non-violent but instructional videos to YouTube, or simply using the black flag of ISIS as their avatar/image on any of these social media? Moreover, if we can agree that such acts deserve constraint, given that we are talking about privately owned media companies, whose responsibility is it to constrain these acts? That is, granting that free speech has its limits and that certain terrorist-related speech acts online can be constrained is really only the start of the discussion.

3. THREE ASPECTS: BELIEF, SPEECH, PUBLIC COMMUNICATION

For a more sophisticated analysis, we need to mark out three particular concepts around speech that have particular and distinct moral valences: the freedom of belief, the right to free speech, and the freedom of public communication.

3.1 Freedom of Belief

A challenge for modern liberal democracies is that there is a commitment to the idea that people are free to believe what they will. Liberal democracies recognize that their citizens will have differing conceptions of the good: they are, at their core, pluralistic.

> [R]adical beliefs and extremist attitudes are not necessarily illegal, nor are they inherently negative…It is not altogether uncommon for several individuals to… hold views or opinions that may be considered extreme. In a majority of these cases, violence or any other problematic manifestations of these beliefs will not occur. (Macnair and Frank 2017, p. 149)

A person in a liberal democracy can think whatever they want, including terrorists. What matters is not the belief, per se, but how the belief plays a role in the person's behaviour (Henschke 2017, pp. 213–14). In a liberal democratic society, the state has no business in my beliefs, terrorist or otherwise. That said, simply because someone has a freedom of their beliefs, it does not follow that they should be free from criticism. Any of the moral, social, or political arguments that one brings out in favour of freedom of belief equally apply to criticisms of that belief (Henschke 2017, p. 214).

3.2 Freedom to Speak

When considering free speech, it is the expression of those beliefs that is under review, rather than the beliefs themselves. The point of difference here is when a private set of beliefs is made public. This draws from the basic recognition that a speech act involves a speaker and an audience. 'Speech is plainly not a self-regarding act' (Schauer 1982, p. 10). What interests us here is how speech can be limited – what measures or sanctions can be used to limit a person's speech acts.

> Such sanctions take two major forms. The first, and most serious, is legal punishment by the state, which usually consists of a financial penalty, but can stretch to imprisonment (which then, of course, further restricts the persons free speech). The second threat of sanction comes from social disapprobation. People will often

> refrain from making public statements because they fear the ridicule and moral outrage of others. (van Mill 2017)

When considering terrorism online, liberal democracies have become increasingly active in using state power – or the threat of state power – to constrain certain speech acts online. This has been done by a series of legal means in which a person can be charged with glorifying terrorism, possessing and distributing terrorist material, and so on (Duffy and Pitcher 2018). Drawing from the earlier section, the moral foundations draw from a pluralistic combination of the ideas that acts of glorification and possession of material do not show proper respect to the victims of terrorism, pose potential harms of radicalization, and increase capacity to engage in terrorist acts, and in times of high tension, a need to contain potentially inflammatory material.

When considering social sanction, we need to consider proportionality. As Jon Ronson describes, social sanction online can be vastly disproportionate to an offensive speech act (Ronson 2015, pp. 84–6). If a speaker is to advocate support for a particular terrorist act, it is possible for that speaker's identity to be made public through 'doxxing'; once this is done, the speaker may lose their job and suffer extreme and pervasive harassment. The concern is whether the speaker is actually an active supporter of terrorism – given the easy way that speech acts online can lose their intended meaning, what level of certainty is needed to know that the speaker is actually an active supporter of terrorism? Further, this social sanctioning is a form of mob justice (Ronson 2015, pp. 85, 262). If a key advance of liberal democracies is the development of fair criminal justice systems, then is it fair for mob justice to persist here? The difficulty is in assessing what the state's responsibility is in these situations – if there is an active online mob causing significant and disproportionate harm to a speaker, then does the state have a responsibility to intervene to either protect the original speaker and/or to constrain the virtual mob?

3.3 Free Public Communication

Moving from freedom of belief through freedom of speech, perhaps it is not the speech act that warrants constraint – instead, we might be concerned with the right to *publicly communicat*e that speech act. That is, a person might have a right to free and unconstrained speech, but this does not necessarily mean that they can distribute, publicize, or broadcast what is said. A speech act and the public communication of that speech act are not necessarily equivalent.

First, there is a conceptual distinction between a speech act and public communication. Public communication involves how that speech act is connected to an audience, and the size of that audience. Communication here can range from a speaker in a conversation with a single listener, to the speaker yelling

in a town square to passers-by and news media recording and broadcasting what the speaker was yelling on the nightly news, to friends of the speaker recording it on their smartphones and uploading the audio and video to Twitter, Periscope, YouTube, Facebook, and so on. This communication might occur instantaneously or at some later date. *Public* communication necessarily involves communication that involves an audience that extends beyond interpersonal communication. Thus, speech and public communication are related but different.

Recognizing the distinction between speech and public communication is important as it is inaccurate to say that interfering in a public communication is morally equivalent to a violation of free speech: a speaker cannot walk into a newsroom, demand to have their speech act broadcast immediately, and then complain that their fundamental human rights are being violated if the producers refuse their demands. Just as there are legitimate constraints on free speech, simply because speech and public communication are logically connected, it does not follow that an act of public communication entails unfettered access to whatever means of communication the speaker wants.

4. WHAT TERRORIST ACTS SHOULD WE CONSTRAIN ONLINE?

In liberal democracies, freedom of speech, including the freedom of public communication, is still the default option. However, we recognize that there are situations where such speech acts might legitimately be interfered with. However, to do this, we must first have a way of identifying terrorist speech acts online that warrant intervention. First is the significance of the public communication, in relation to offence/respect, potential harms, and given the immediate context in which the speech act is occurring. Second is the public interest. That is, there may be a case for intervening in a terrorist communication, but this could be outweighed by the public interest.

Significance is vague and open to interpretation, and draws from subjective experience and the historical factors noted in the discussion of hate speech. We also have to consider harder cases that involve less significant communications whose glorification is controversial – a person posting the black flag of ISIS to their Facebook feed and so on. These are 'glorification' in some soft sense, but hardly rise to the level of significance. For these, it would seem that the prima facie case for free speech holds. Given the default protections for free speech in liberal democracies, the significance has to actually be *significant*. Even in times of high stress and social tension, where the bar for significance is justifiably lowered, such lowering would be minimal and, most important, temporary.

A further point on this is that any active interference in a speech act needs to be accountable. That is, given that the default setting is non-interference, and potentially active protection of a right to free speech, interference with the public communication needs to be a reflective and explicable set of explanations. Such explanation would *start* with the claim to significance, and then go into detail about what is meant by significance in this situation – who is being particularly disrespected by the speech act, how likely it is to cause harm, how significant are those harms, who is harmed by the unfettered speech act, how and why does this context/situation require intervention, what conditions would obtain for the freedom to speak to return, and whose rights are being interfered with? In response to IS's actions and the ensuing public criticism, sites like YouTube and Twitter have developed and sought to enforce their terms of use, being more active in taking down offensive messages and content, closing accounts, and banning particular people from setting up new accounts. However, they face problems of explicability and accountability around the implementation of these policies.

5. RESPONSIBILITY AND TERRORIST SPEECH ACTS

To answer questions about who holds responsibilities for constraining public communications of terrorist speech acts, we need clarity on the rights of public communication. Assuming that there is some basis to a claim about free public communication – is such freedom a negative right or a positive one? In Isaiah Berlin's classic formulation (Berlin 1959), a negative right involves a freedom from interference. If a speaker has a claim for a negative right of public communication, then no one (without good reason at least) has a right to *stop* that public communication. In contrast, a positive right involves a claim that others need to *support* the given act. If a speaker has a claim for a positive right of public communication, then others should actively help them in their communication.

To bring this to public communication and terrorism online, in terms of negative freedom, while people might have a prima facie claim to non-interference in their communications, the state has a set of legal, moral, and social responsibilities that can justify interference in speech acts that show significant disrespect for individuals, can cause significant harm, and/or occur at times of significant social unrest. Here, posting a beheading video online, distributing how-to manuals, and so on could potentially warrant active interference in that public communication.

While less significant speech acts might maintain freedom from interference and a speaker can say whatever they want using whatever means of public communication they have access to, they do not necessarily have a claim to access all means of public communication. For instance, a speaker might be

entitled to set up a website where they can publicly communicate whatever speech acts they want. However, to demand that Facebook, Twitter, or the like must *broadcast* those speech acts is making a demand about *positive* freedom. Following Berlin, this claim about positive freedom generally requires a much stronger set of justifications.

Things become more complicated when considering whether social media have a positive duty to allow terrorists and other extremists to publicly communicate using their platforms. It would seem here that, as these companies are privately run enterprises, they have no special responsibility to ensure a *positive* freedom of communication. However, this is becoming a more controversial claim. Facebook, Twitter et al. are evolving. While they were originally simply 'social networks' – ways for individuals to stay in contact with each other – their reach and impact into our personal, social, and political worlds is extensive and powerful. If we hold that their moral authority and accompanying responsibilities derive from what sort of institution they are (Aradau 2010), I suggest that our assessment of their moral responsibility turns on whether we consider them social equivalents to news media, or public infrastructure.

For instance, given their reach and impact, they have evolved to become comparable to news media. Moreover, insofar as news media and journalism play a particularly important moral role as social institutions, social media inherit the moral responsibilities of news media and journalism. This cuts two ways: news media are typically granted particular freedoms about public communication if that news is in the public interest, but they have a set of responsibilities that come with their capacity to broadcast public communications to a wide audience.

> The liberty of the press is indeed essential to the nature of a free State…Every freeman has an undoubted right to lay what sentiments he pleases before the public: to forbid this, is to destroy the freedom of the press: but if he publishes what is improper, mischievous, or illegal, he must take the consequence of his own temerity. (Blackstone, quoted in Wacks 1995, p. 32)

Journalists and 'news media' as an institution have a set of ethical and professional responsibilities, again, tied to the notion of public interest. As we have discussed, speech acts can justifiably be constrained if they are offensive, harmful, or expressed in a particular context. These conditions clearly relate to the public interest; if we consider social media to be morally equivalent to news media, then, like news media, public communications using these platforms need also be constrained when such communications are detrimental to the public interest.

As argued elsewhere, if social media 'are, instead, something more like public infrastructure – like that of a road system or energy system – then

they may have to constrain their responsibility to shareholders and profits by reference to public safety' (Henschke and Reed, 2021). Drawing from a principle like public safety, we can return again to the earlier discussion of the constraints on free speech. In particular, the constraint on speech acts that can cause significant harm is directly applicable to public safety. Thus, if social media are considered social infrastructure, then public communications that threaten public safety can potentially be constrained. That said, historically, protections for free speech, such as the US First Amendment, de-emphasize:

> the link between speakers and the harms speech can cause…[S]peakers are not held responsible…for entirely foreseeable criminal actions they inspire third parties to take…This stands in sharp contrast, for example, to product liability cases, in which we hold manufacturers responsible for injuries they should have foreseen and could have prevented, even if those injuries would not have occurred but for someone else's subsequent negligence or criminality. (Weinberg 1993, p. 1145)

On this account, if social media are a public infrastructure, like those manufacturers that produce goods, what matters is public safety.

The overall point here is, if we now consider social media to be something like news media or public infrastructure, there are reasons to consider that they have a responsibility to constrain public communications of terrorist material if those public communications cause significant offence, harm, or occur in contexts of heightened security concerns. On the one hand, if social media are considered to be evolving into something like news media, then, while social media might have particular freedoms associated with a free press, they also have responsibilities that align with those of the news media. Second, if, instead, we see social media as something more like providing a fundamental social good, they are more like infrastructure. This, then, carries with it a responsibility to public safety. Either way, the evolution of social media brings with it special moral responsibilities about online communications that may justify constraints on speech.

6. CONCLUSION

As social media have evolved into essential features of our social and personal lives, we are faced with the conflict between ideals, such as free speech, and the threats posed by malicious actors, such as terrorists. This presents a perennial challenge: on the one hand, we need and hope for governments to discharge their responsibility to ensure our security. On the other hand, 'governmental control over editorial policies typically will be exercised in a discriminatory fashion, privileging that which is in vogue, mainstream, and safe while handicapping that which is not' (Krattenmaker and Powe 1995, p. 1733). This worry about editorial control is arguably even greater when considering institutions

such as social media that are not only relatively inexperienced in exercising editorial control, but also frequently opaque and at times capricious in how their editorial control is exercised. The public communication of terrorist speech acts online has exposed how necessary it is for something to be done to constrain such acts. However, it also reminds us of how complex the interface between the principles of free speech, the need for security, and the responsibilities of traditional and new media is. This chapter does not seek to answer questions of who ought to do what in situations of terrorism online. Instead, it is an effort to expose the moral mechanics of this interface in the hope that we can get some better sense of how to look for answers to these challenges.

NOTES

1. Frederick Schauer describes Milton's *Areopagitica* as 'the earliest comprehensive defence of freedom of speech' (Schauer 1982, p. 15).
2. There's a large and comprehensive body of literature on this – Sunstein's (1993) *Democracy and the Problem of Free Speech*, Sadurski's (1999) *Freedom of Speech and Its Limits*, and Schauer's (1982) *Free Speech* are all good places to start.
3. The US, for instance, permits public displays of symbols like the swastika, while Germany holds such speech acts to be illegal and punishable. Social and legal approaches to free speech vary across liberal democracies.
4. See other chapters in this collection for more on this and related issues.
5. A violent act, like a suicide bomb attack, may be a communicative act (but not a speech act, as it is not a linguistic act). However, this communicative aspect means that there is a non-violent aspect essential to the act. For a discussion of these issues, see Tony Coady's article, 'The Idea of Violence' (1986).
6. I do not want to enter into discussions of what counts as 'child pornography'. For instance, the notion of what counts as 'pornography' – 'I'll know it when I see it' – is an open and essentially contested concept. The point is to use whatever notion a person considers to be 'child pornography' as a moral foundation for something that almost all of us would consider not to be protected by claims of free speech. That is, to say, 'I can produce and/or distribute and/or view child pornography *because* of a right to free speech and free speech alone', is an argument most people would reject due to the moral problem of child pornography.
7. See also Feinberg's *Offense to Others* (1985) for more on this.
8. As will be discussed, the constraints are not, perhaps, on belief or even speech. Rather, when considering the focus on the online environment, we might need to think of constraints on *public* communication.

REFERENCES

ABC News (2016), 'Islamic State Call For Attacks on Specific Locations in Australia "Propaganda", Victorian Police Say'. *ABC News*, accessed 7 September 2016 at http://www.abc.net.au/news/2016-09-06/is-calls-for-attacks-in-australia-dismissed-as-propaganda/7819774.

Aradau, Claudia (2010), 'Security That Matters: Critical Infrastructure and Objects of Protection', *Security Dialogue* 41 (5), 491–514. doi: 10.1177/0967010610382687.
Awan, Imran (2017), 'Cyber-Extremism: Isis and the Power of Social Media', *Society* 54 (2), 138–49. doi: 10.1007/s12115-017-0114-0.
Berlin, Isaiah (1959), *Two Concepts of Liberty: An Inaugural Lecture Delivered before the University of Oxford on 31 October 1958*, Vol. 31, Oxford: Clarendon.
Coady, Tony (1986), 'The Idea of Violence', *Journal of Applied Philosophy* 3 (1), 3–19.
Duffy, Helen, and Kate Pitcher (2018), 'Inciting Terrorism? Crimes of Expression and the Limits of the Law', in Benjamin J. Goold and Liora Lazarus (eds), *Security and Human Rights*, London: Hart Publishing.
Feinberg, Joel (1985), *Offense to Others*, London: Oxford University Press.
Fish, Stanley (1994), *There's No Such Thing as Free Speech: And It's a Good Thing, Too*, Oxford: Oxford University Press.
Friedman, Thomas L. (2005), *The World is Flat: The Globalized World in the Twenty-First Century*, London: Penguin.
Gates, Scott, and Sukanya Podder (2015), 'Social Media, Recruitment, Allegiance and the Islamic State', *Perspectives on Terrorism* 9 (4), 107–16.
Griffin, James (1986), *Well-Being: Its Meaning, Measurement and Moral Importance*, Oxford: Oxford University Press.
Henschke, Adam (2017), *Ethics in an Age of Surveillance: Virtual Identities and Personal Information*, New York: Cambridge University Press.
Henschke, Adam, and Alastair Reed (2021), 'Toward an Ethical Framework for Countering Extremist Propaganda Online', *Studies in Conflict & Terrorism*. doi: 10.1080/1057610X.2020.18667.
Kilcullen, David (2016), *Blood Year: The Unravelling of Western Counterterrorism*, Oxford: Oxford University Press.
Klausen, Jytte (2015), 'Tweeting the Jihad: Social Media Networks of Western Foreign Fighters in Syria and Iraq', *Studies in Conflict & Terrorism* 38 (1), 1–22. doi: 10.1080/1057610X.2014.974948.
Krattenmaker, Thomas, and Lucas A. Powe, Jr. (1995), 'Converging First Amendment Principles for Converging Communications Media', *The Yale Law Journal* 104 (7), 1719–41.
Laqueur, Walter (1977), *A History of Terrorism*, New Brunswick: Transaction Publishers.
Larmore, Charles (1987), *Patterns of Moral Complexity*, Cambridge: Cambridge University Press.
Leetaru, Kalev (2018), 'Is Twitter Really Censoring Free Speech?', *Forbes*, 12 January 2018, accessed 5 April 2021 at https://www.forbes.com/sites/kalevleetaru/2018/01/12/is–twitter-really-censoring-free-speech/#3b3416ea65f5.
Macnair, Logan, and Richard Frank (2017), 'Voices against Extremism: A Case Study of a Community-Based CVE Counter-Narrative Campaign', *Journal For Deradicalization* 10, 147–74.
Morozov, Evgeny (2013), *To Save Everything Click Here: Technology, Solutionism and the Urge to Fix Problems that Don't Exist*, London: Penguin.
Qasir, Sophia (2013), 'Anonymity in Cyberspace: Judicial and Legislative Regulations', *Fordham Law Review*, 81 (6), 3651–92.
Ronson, Jon (2015), *So You've Been Publicly Shamed*, London: Picador.
Sadurski, Wojciech (1999), *Freedom of Speech and Its Limits*, Dordrecht, Boston: Kluwer Academic Publishers.

Schauer, Frederick (1982), *Free Speech: A Philosophical Enquiry*, Cambridge: Cambridge University Press.
Schenck v. United States (1919), 249 U.S. 47, 39 S. Ct. 247, 63 L. Ed. 470.
Shehabat, Ahmad (2012), 'The Social Media Cyber-War: The Unfolding Events in the Syrian Revolution 2011', *Global Media Journal: Australian Edition* 6 (2), accessed 5 April 2021 at http://www.hca.westernsydney.edu.au/gmjau/archive/v6_2012_2/ahmad_shehabat%20_RA.html.
Sunstein, Cass R. (1993), *Democracy and the Problem of Free Speech*, New York: The Free Press.
van Mill, David (2017), 'Freedom of Speech', in Edward N. Zalta (ed), *The Stanford Encyclopedia of Philosophy*, accessed 20 October 2017 at http://plato.stanford.edu/archives/win2013/entries/freedom-speech/.
Wacks, Raymond (1995), *Privacy and Press Freedom*, London: Blackstone Press.
Waldron, Jeremy (2009), '2009 Oliver Wendel Holmes Lectures: Dignity And Defamation the Visibility of Hate', *Harvard Law Review* 123 (7), 1596–657.
Weckert, John (2007), 'Giving and Taking Offence in a Global Context', *International Journal of Technology and Human Interaction* 3 (3), 25–35.
Weinberg, Jonathan (1993), 'Broadcasting and Speech', *California Law Review*, 81 (5), 1103–206.

11. Counter-terrorism and PSYOP

Michael Robillard

1. INTRODUCTION

Standard thinking about the ethics of war is often cashed out in familiar terms of '*jus ad bellum*' and '*jus in bello*'. '*Jus ad bellum*', or the ethics of going to war, typically articulates three morally justifiable reasons for a sovereign state to wage war. These reasons include: response to aggression, pre-emptive response to imminent threat, and humanitarian intervention to prevent atrocity or crimes against humanity. *Jus in bello*, or the ethics of fighting well in war, typically concerns itself with principles aimed at ethically restraining combatant behaviour once soldiers are already in battle. These principles include adherence to necessary and proportionate harming between legitimate combatants and the principle of non-combatant immunity.[1] Both *ad bellum* and *in bello* principles find explicit articulation in present-day Geneva Conventions, international humanitarian law (IHL), and soldier Rules of Engagement. Noticeable in both *jus ad bellum* and *jus in bello* criteria is their exclusive focus upon *kinetic* actions and *kinetic* states of affairs. Accordingly, for standard *ad bellum* criteria to find purchase, some sort of physical or territorial boundary must be crossed, or about to be crossed, or some physical atrocity must be occurring or about to occur. Similarly, for standard *jus in bello* criteria to find purchase, some sort of kinetic exchange between combatants is presumed to be occurring or about to occur.

With the advent of the 21st-century informational age, the increased use of informational warfare and PSYOP (Psychological Operations) methods has significantly problematized standard *ad bellum* as well as *in bello* thinking. Indeed, without a physical territorial boundary being overtly crossed, without a designated battlefield, and without physical infrastructure or human bodies taking noticeable and direct kinetic damage, it is unclear how exactly informational warfare and PSYOP means and methods fit within the contemporary just-war paradigm, if at all.

On one hand, since PSYOP fail to rise to the level of a discernible kinetic harm, it can be argued that any and all non-kinetic and/or informational methods of shaping, deception, espionage, and subterfuge should be fair game

amongst state actors and should not be counted within *ad bellum* or *in bello* considerations whatsoever. Following this reasoning, strict just-war readings of the concept of *necessity* would dictate that use of informational or PSYOP methods would always be preferable to the use of kinetic options since informational methods would presumably always be less physically harmful and less lethal.

This would then leave open methods of election hacking, hacking and hacktivism more broadly, forms of lawfare, mass trolling and mass disinformation campaigns, propaganda and false-flag efforts, lawfare, institutional subversion, weaponized protests and marches, astroturfing, and weaponized NGOs, among other soft-war methods as completely fair game between competing nation states (see Gross and Meisels 2017).

On the other hand, it can be argued that informational and PSYOP methods ought to be considered within the purview of *ad bellum* as well as *in bello* considerations for it is both a sociological and pragmatic fact that modern-day militaries, to include the US military and its allies, *do* indeed conceive of the practical scope of 'war' as encompassing an informational domain prior to and below the level of armed conflict as well as throughout the phases of armed conflict (see Scharre 2016). Consequently, since modern militaries conceive of war in these terms, it is incumbent upon just-war theorists to count informational warfare and PSYOP as theoretically and practically relevant within their scope of concerns. Here, Rawls's principle of reflective equilibrium pulls us in the direction of *praxis* over theory.[2] What is more, if we extend our conceptions of *necessity*, *proportionality*, *collateral damage*, and *harm* to include downstream second- and third-order effects resulting from informational and PSYOP campaigns, then the use of such non-kinetic methods could conceivably be *more* harmful in the aggregate than the use of standard kinetic options (both in virtue of direct psychological harms as well as second-order kinetic harms). Indeed, arguments for the moral worseness of economic sanctions versus traditional military campaigns seem to appeal to these same intuitions.[3] Accordingly, such forward-looking calculations therefore skew the commonplace 20th-century intuition that non-kinetic options are always ethically preferred over direct kinetic options.[4]

Problematizing matters further is consideration of what ethical factors and reasons inform and constrain the use of informational and PSYOP methods when it comes to *counter-terrorism and counter-insurgency specifically*. Indeed, conventional warfare and terrorism/counter-terrorism differ radically in many important respects. It is therefore highly controversial if/how standard just-war principles should apply to a counter-terrorism context. At the very least, it seems that we need to adapt such principles such that they are sensitive to the nuances and subtleties of such irregular contexts, for even if we did discern a clear set of normative principles that could tell us what necessary

and proportional PSYOP means and methods looked like between legitimately recognized nation states in a conventional context, the wild card of terrorist organizations and illegitimate sub-state actors operating within the domain of another legitimate nation state's physical territory ostensibly adds a new set of variables within our overall ethical calculus.

In Section 2 of this chapter, I give an explanation of various contemporary PSYOP means and methods. In Section 3, I articulate where certain ethical values, tensions, and trade-offs arise with the employment of these methods generally. In this section, I also explain how these ethical considerations specifically apply within the counter-terrorism operational space. Borrowing recent concepts from LTC Bob Underwood, this chapter emphasizes two main reciprocal ideas. First, *communicative acts* often have normatively laden kinetic effects. In other words, PSYOP and informational operations will often cause foreseeable as well as unforeseeable second-order kinetic effects with ethical valence. For instance, a mass disinformation campaign about a political coup could cause mass riots and infrastructural damage in the streets as well as deep suspicion and distrust of social and epistemic institutions for months or even years to come. Second, kinetic operations will often generate foreseeable and unforeseeable second-order communicative effects (see Underwood 2019). For instance, a kinetic drone strike could create a narrative effect that could lose the hearts and minds of the local populace. Accordingly, thinking about ethical and efficacious counter-terrorism operations ought to regard these communicative components (as both causes and effects) as one if its highest strategic priorities.

2. PSYOP: NATURE, MEANS, AND METHODS

According to 'Joint Publication 3-53 Doctrine for Joint Psychological Operations', 'Military PSYOP constitutes a systematic process of conveying messages to selected foreign groups to promote particular themes that result in desired foreign attitudes and behaviours. PSYOP are used to establish and reinforce foreign perceptions of US military, political, and economic power and resolve' (Department of the Army and Department of the Navy 2003). While non-kinetic and informational means have been used throughout history by political and military leaders to influence foreign militaries and adversaries, contemporary PSYOP campaigns in the informational age aim to communicate directly and with much greater precision and fine-grained accuracy to a variety of intended targets and audiences across a variety of mediums and with a variety of messages (Department of the Army and Department of the Navy 2003). These mediums often include Internet and social media, television and radio, art, literature, print journalism, and pamphlet campaigns. Messaging is often directed at and tailored to a variety of audiences, to include key enemy

leadership, mid-tier officers, foot soldiers, allies of enemies, and civilians. Typical messaging frequently includes content involving promises or threats of force or retaliation, conditions of surrender, safe passage for deserters, invitations to sabotage, support to resistance groups, as well as other messaging. New communication technologies have also opened up the increased capacity for military units to 'narrowcast' a given narrative or message to a pinpointed person or niche group.

PSYOP missions in a military context frequently involve such core actions as:

1. Influencing foreign populations by expressing information subjectively to influence attitudes and behaviour and to obtain compliance and non-interference. This type of information can facilitate military operations, minimize needless loss of life and collateral damage, and further the objectives of the United States and its allies.
2. Advising the supported commander through the targeting process, regarding Psychological Actions (PSYACT), PSYOP-enabling actions, and targeting restrictions to be executed by the military force.
3. Serving as the supported military commander's voice to foreign populations to convey intent.
4. Countering enemy propaganda, misinformation, and opposing information to correctly portray friendly intent and actions, while denying others the ability to polarize public opinion and political will against the US and its allies within an area of operations (Department of the Army and Department of the Navy 2003).

Given such 21st-century informational thinking, to conceive of warfare exclusively in terms of kinetic damage and what philosophers refer to as eliminative harming (that is, the physical killing or removal of the enemy from battle) completely ignores the communicative elements inherent in such acts along with the ethical and prudential elements endemic to them. Thus, to strategically bracket off PSYOP from kinetic warfare further reinforces this false dichotomy. The fact that these kinetic actions are frequently also used to perform communicative actions is often ignored. Moreover, these communicative actions frequently have an ethical dimension that cannot be reduced to the ethical dimension of the kinetic actions that are their means of communication.

As LTC Bob Underwood notes in a recent essay:

> Killing in war eliminates threats but also plays a part in influencing the decisions of other persons beyond those we might kill. This suggests that killing in war has a communicative function, and that the message is an important consideration that can feature in the balance of reasons to kill some but not others in war. This is true provided combatants can permissibly kill some as means to communicate to others.

> I argue that just combatants, those that fight for just aims, can permissibly kill to communicate and that unjust combatants cannot. This is a new reason to revise our intuition that combatants on both sides hold equal rights to kill, the so-called moral equality of combatants (MEC). (Underwood 2019)

Vindication of these points can be most readily seen in Harry Summers's (1982) famous work, *On Strategy: A Critical Analysis of the Vietnam War*. Here, Summers notes that US strategic thinking in Vietnam was fundamentally flawed insofar as it understood military victory through a predominantly Clausewitzian lens and saw military success as amounting to a mere set of individual aggregated kinetic victories. Throughout the entirety of the Vietnam War, the US failed to ever lose a single kinetic engagement, though they ultimately lost the war. Hence, on Summers's analysis, in overly focusing on securing positional and kinetic victories, the US military failed to 'win the hearts and minds' of both the Vietnamese people in theatre as well as the American public back home. Failure to properly acknowledge the communicative or narrative elements of kinetic actions in war ostensibly functioned as one of the major reasons the US war in Vietnam was ultimately lost. Insofar as we regard quick military victory as being morally superior to protracted warfare and excessive bloodshed, then the US's failure to properly recognize the communicative elements inherent in their kinetic operations arguably also carried with it a significant moral, and not just prudential, component.

Turning our attention now to the flip side of the coin, just as kinetic actions generate intended or unintended communicative effects, non-kinetic PSYOP can and likely will generate second- and third-order indirect causal effects that could likely result in kinetic damage or kinetic harms, intended or unintended.[5] Such a connection is at least recognized in domestic law. Indeed, an average person ostensibly has First-Amendment rights of free speech, but those rights are curtailed by considerations of public endangerment, hence, the prohibition on yelling 'fire' in a crowded movie theatre. Analogously, it can be argued that, despite being non-kinetic in nature, PSYOP shaping operations (that is, propaganda campaigns, public demonstrations, protests, lawfare, and so on) could plausibly entail a set of predictable downstream kinetic harms (for example, riots, social discord, mass panic, revolution, and so on) equal to or greater than actual kinetic attacks. Setting such predictable causal chains into effect would ostensibly carry with it not just a strategic but also a moral component. Even if one does not set such predictable causal chains into effect via PSYOP shaping operations, then arguably one is still doing something morally negligent by facilitating such preconditions for downstream harm.

The main upshot from the following analysis is acknowledgement that the narrative or communicative component of warfare, resulting from non-kinetic PSYOP actions or kinetic actions, *significantly* matters within both ethical

and prudential thinking about warfare. Since the public, the international community, members of the military, coalition partners, terrorist enemies, and civilian non-combatants will all ostensibly be interpreting certain kinetic effects, then the narrative management of all of these audiences *simultaneously* must be considered and, indeed, prioritized. As a result of the Internet and the information age, kinetic actions and PSYOP actions no longer occur within an informational vacuum contained to a specific geographic region of physical conflict or battlefield.

To be fair, present US military doctrine at least gives cursory acknowledgement to the informational or communicative attributes of kinetic actions. Indeed, JP 3-53 states:

> It is important not to confuse *psychological impact* with PSYOP. Actions such as shows of force or limited strikes may have a psychological impact, but they are not PSYOP unless the primary purpose is to influence the emotions, motives, objective reasoning, decision making, or behavior of the foreign target audience. (Department of the Army and Department of the Navy 2003)

JP 3-53 continues:

> [The] very activity of the force has psychological implications that may be leveraged in the battle to influence a foreign TA [target audience]. If communicated to the potential opponent, such things as the arrival of the force in the operational area, the multinational nature of the force, its combat power, technological sophistication, level of training, and preparation of US and multinational forces can break the adversary's will to fight. (Department of the Army and Department of the Navy 2003)

While contemporary US doctrine does, in fact, technically acknowledge this intertwinement of kinetic and informational/communicative features, I argue that, both pragmatically and ethically, the importance of this connection has been undervalued. Thus, management of these communicative elements within strategic thinking not only has significant ethical and prudential bearing in general but also has specific import within the counter-terrorism space, as we will now see.

3. APPLICATION TO COUNTER-TERRORISM

Before delving into the specifics of the ethics of counter-terrorism, I must make the disclaimer that none of the normative claims here should be understood as *in principle* claims that hold necessarily in all possible worlds. Rather, they should be better understood as prima facie ethical rules of thumb that could conceivably be overridden given real-world operational constraints and contingencies. Hence, in some plausible scenarios, it could

be an all-things-considered good to stick to PSYOP and non-kinetic informational pressures to achieve a certain set of real-world effects. In other instances, it could be ethically preferable to 'go kinetic' and to resort to limited kinetic targeting, reprisals, or decapitation efforts. Lastly, though counter-intuitive, there could still be plausible scenarios where the overall set of predictable goods to be achieved or predictable harms to be averted would justify or even obligate military leaders to opt for a large-scale kinetic effort over a non-kinetic PSYOP method. All this being said, I nonetheless argue that there are weighty and significant ethical reasons connected to the communicative features of both kinetic and PSYOP efforts within the counter-terrorism space that warrant much greater attention.

To put the above claim in more concrete terms, consider both the strategic and ethical implications of the communicative effects connected to the now infamous photos from Abu Ghraib. Had the events at Abu Ghraib happened in an informational vacuum, say 100 years ago, the strategic and normative import, though still heinous, would have been contained to that one isolated event (see Johnson et al. 2016). Now, with the age of the Internet and the velocity of information, those few images of ethical misconduct likely did more for al-Qaeda (AQ) recruitment than any AQ-led recruitment effort before or since, and likely resulted in a much greater loss of strategic resources, time, money, and most importantly, lives. The take-home point from such an event is the recognition that isolated kinetic behaviours can have drastically deleterious communicative, and therefore strategic, effects.[6]

Similar communicative elements are endemic not just to morally salient areas within counter-terrorism, such as prisoner detention, but also to the area of drone strikes and targeted killing. From a strategic lens, HVTs (high-value targets) are often evaluated networks, and have a known likelihood of creating immediate or future harm.[7] From an ethical lens, as Underwood (2019) notes, military ethicists often cash out conceptions of necessary and proportionate harming of culpable targets in primarily *eliminative* terms. Thus, predominant strategic and ethical thinking both drastically undervalue the *communicative* effects, intended or unintended, intrinsic to such targeting acts.

The communicative effects of targeted killing efforts trace along a variety of axes and require the consideration of multiple audiences simultaneously. For instance, all things being equal and occurring within a hypothetical vacuum, it might turn out that it is both strategically beneficial and ethically justified to perform a targeted strike on a culpable member of a terrorist network. However, once we place the targeted strike within a real-world context and consider how the communicative effects of such a targeted strike (and its visual aftermath) will likely be perceived by various audiences (the local civilian community, the larger Islamic community, AQ leadership,

middlemen, foot soldiers, the host nation's leadership, the world community at large, American soldiers, Amnesty International, the Red Cross, Human Rights Watch, and the American public back home), then such a targeted strike on a culpable aggressor might no longer be ethically justified or even strategically prudent (Ludvigsen 2018). Hence, it is not enough that, from the point of view of the Geneva Conventions or IHL standards, a given target meets the criteria for justified targeting. Indeed, not only must such a justification obtain, but it also must be the case that the various audiences noted positively understand and recognize such justifications, if only partially. In other words, ethical and effective PSYOP might include predisposing a given target populace not only to certain facts but also to certain moral attitudes assuming that that populace is likely to act (for example, protest the war, join the enemy) or desist from acting (for example, refuse to be conscripted) on the basis of certain understandings.[8]

Accordingly, it is both strategically and ethically preferable that any targeted killing decision be complimented with well-thought-out PSYOP and informational efforts to effectively communicate such justifications to the various audiences mentioned. Otherwise, absent such a communicative effort, targeted killing acts, especially ones incurring collateral damage, run the very real hazard of becoming little more than propagandistic fodder for terrorist networks to leverage to increase their recruitment numbers and strategic messaging. Short of this, such acts still run the risk of being interpreted by the local and world community as yet another instance of unnecessary bloodshed. Lastly, there is arguably a slow and sustained psychological and communicative effect of the *potential* of a targeted strike at any moment that arguably has ethical and strategic relevance to a civilian population under a sustained drone campaign. Accordingly, military ethicists and military strategists alike should take more seriously the communicative effects laden in each kinetic act and how such effects will likely be interpreted by various audiences and stakeholders.[9]

Now that we have looked at some of the ethical implications of kinetic actions with intrinsic communicative or psychological effects, let us turn to consider the converse – that is, non-kinetic or PSYOP efforts that have intrinsic or downstream kinetic effects. As stated briefly in the introduction to the chapter, one could argue that PSYOP and non-kinetic methods of warfare are *always* morally preferable to instances of kinetic harming. I, however, disagree with this claim and argue that certain PSYOP or non-kinetic informational efforts could conceivably be morally *worse* than kinetic acts in certain cases (because of resulting kinetic or even mere attitudinal effects). For instance, a populace who has been conditioned via PSYOP to be in perpetual fear of imminent drone targeting at any instant could arguably be said to be worse off than a similar populace who experiences short-term kinetic

collateral damage via tanks and troops, but who nonetheless psychologically regards the military occupation to be effectively over.

As noted previously, one instance of this is the domestic case of yelling 'fire' in a crowded building. While the utterance of the sound 'fire' has no intrinsically harmful features to it, doing so in the social context of a movie theatre could nonetheless generate the predictable harm of mass trampling. Taken to an operational context on a macro-scale, certain PSYOP efforts could set off predictable or even unpredictable social effects eventually resulting in the downstream result of riots, social upheaval, and revolution that are much greater in aggregated harm, much more indiscriminate, and much less controllable than a limited drone strike or reprisal via artillery shelling. Once again, just as kinetic acts will have intended or unintended communicative effects to various audiences, the downstream kinetic effects of PSYOP efforts so too will carry with them intrinsic communicative effects. In real-world terms, a PSYOP effort that inadvertently snowballs into a bloody riot might generate similar bad optics as a smouldering crater from a recent drone strike. Hence, PSYOP or informational efforts within the counter-terrorism space must therefore take into account not only the immediate ethical and prudential implications of the given PSYOP act but also the ethical and prudential implications of the optics and communicative effects of a PSYOP effort's predictable downstream consequences.

What is more, even if a given PSYOP effort did not result in any deleterious downstream kinetic consequences or negative communicative effects, there are still arguably situations where non-kinetic PSYOP efforts could still conceivably be worse than kinetic efforts given their attitudinal effects on a population. Indeed, if a military force aimed to radically subvert every major institution of a given populace or muddied the waters of a populace's shared epistemic space (that is, through the use of grey or black propaganda, spin-doctoring kinetic actions, leveraging selective facts to achieve certain social effects, and so on) so drastically that all social trust and social cohesion broke down, then such epistemic harms, though non-kinetic, could arguably count as morally worse than kinetic harming. Intuitively this seems correct since the vast majority of persons, if given the choice between having an arm broken or having their reputation completely and irreparably ruined would likely opt for the former rather than the latter. We might think the same goes, for instance, for collective humiliation or collective reputational disrespect.

A final morally salient area in the counter-terrorism space that is worth noting and relevant to PSYOP effects as well as kinetic actions with psychological/communicative effects is the area of *negotiation*. Since the Reagan era, it has been taken as gospel by the military community that the US and its allies 'never negotiate with terrorists' (see Bapat 2006). If we understand 'negotiation' in a very restricted sense, in terms of sitting down at a peace

table and regarding the terrorist organization as a legitimate political entity on par with them in all respects, this maxim seems to hold true. However, and this is to return to Underwood's earlier claim, PSYOP efforts and kinetic actions both carry communicative effects. Hence, done well, PSYOP and targeted killing efforts against terrorist organizations should not only see their actions as physically eliminating terrorists but also as a means of *communicating* with terrorists. Hence, political violence at the highest strategic level should be seen as a form of signalling ideas, narratives, threats, choices, and reasons. This therefore includes using military methods to counteract active terrorist propaganda. Accordingly, strategic targeting decisions as well as PSYOP efforts should be used in conjunction as a means of communicating certain messaging to members within a given terrorist organization as well as to civilians in a given space. Hence, the core of such messaging, in its most simplistic form, might look something like:

'If you do X, then we will respond with force Y, because of ethical reason Z'.
'You did act X'.
'We now respond with force Y to honour reason Z'.

Importantly, in such efforts, this pairing of kinetic effects and communicative effects would be sensitive to the values set of the audiences most likely to be immediately affected by such acts. In other words, 'ethical reason Z' must actually count as worthy of respect by such intended audiences. Hence, pre-empting and following up a targeted strike with appeals to verses from the Koran that reference codes of conduct for honourable warriors will be arguably more ethical and efficacious than appeals to articles and subclauses taken from contemporary IHL. Similarly, referencing AQ's own stated codes of conduct and communicating how the person targeted failed to meet those standards by AQ's own ideals would arguably be preferable to a PSYOP effort that referenced the target's failure to live up to Westphalian formulations of just conduct. Such appeals to reason could and should also arguably be effectively incorporated into *post bellum* thinking as well. Accordingly, under this wide-scope notion of 'negotiation', terrorist organizations, their supporters, and regional civilians could arguably be communicated with more ethically and effectively while taking the aforementioned reasons into consideration.

4. CONCLUSION

In conclusion, the major take-home message of this chapter is that the counter-terrorism fight is as much of an *informational fight* as it is a kinetic one. Adhering to the antiquated Clausewitzian notion of 'breaking the enemy

army in the field' or the trolley-problem style of ethics both mutually fail to account for the communicative elements of kinetic and non-kinetic acts endemic to war. Accordingly, as we move further into the 21st-century informational age, kinetic and non-kinetic acts in war will become increasingly transparent to the public back home, to the international community, to military members, and to civilians in areas of operations. Given that such informational and communicative effects will be indelibly yoked to military actions, it is incumbent upon military practitioners to ensure that their strategic and operational planning takes into account these factors and how they will likely be perceived by a variety of audiences. That being said, it is not simply enough that counter-terrorism efforts of the kinetic and PSYOP variety operate from a place of heightened ethical sensitivity. Rather, what is needed is that such moral reasoning is not only built into strategic actions, both PSYOP and kinetic, but that such moral reasons are communicated to various audiences *in terms that they will understand and take seriously*. To do so therefore requires that the moral reasoning and the communication of moral reasoning be laden throughout any major operational act, kinetic or otherwise, from beginning, middle, and end.[10]

NOTES

1. For an in-depth overview of contemporary just-war principles, see Lazar and Frowe (2018).
2. For an in-depth explanation of Rawlsian reflective equilibrium, see Daniels (1979).
3. Indeed, it is important to note that economic sanctions can sometimes be more harmful and morally worse than kinetic warfare options. This point at least gives some plausibility as to the moral permissibility of PSYOP methods.
4. Indeed, a sharp and abrupt kinetic action might, in total, be the most humane option as opposed to a long, drawn-out, and indiscriminate PSYOP/soft-war campaign.
5. This is not to suggest that the morality of PSYOP is fully determined simply by the kinetic effects resulting from such acts. Indeed, changing a social group's institutions and epistemic orientation to the world ostensibly carries with it its own moral valence independent of any kinetic harms that may result. My point here is simply to emphasize that the moral evaluation of PSYOP acts, much like kinetic acts in war, should include foreseeable (and possibly unforeseeable) downstream causal chains created by the initial non-kinetic action. The morality of yelling 'fire' in a crowded theatre and lighting a match in a forest full of dry tinder is not contained to the singular act in isolation.
6. Indeed, because of the so-called 'velocity of information' and the networked nature of the Internet, countering enemy narratives in a way that is faster and further reaching is highly difficult once a secret is leaked or a 'bad optics' image is released online and distorted by the enemy. See the United Nations Office of Drugs and Crime publication, 'The Use of the Internet for Terrorist Purposes' (2012).

7. For an in-depth treatment of terrorist propaganda related to US drone operations, see Ludvigsen (2018).
8. Here, we might need to make a distinction between non-manipulative truth-based information campaigns versus propaganda. Indeed, if one is dealing with a recalcitrant or highly epistemically insular populace, then the latter might be a necessary first means in service of the former later on down the road. We might also make the distinction here between 'white', 'grey', and 'black' propaganda. Whereas white propaganda is intentionally attributed to one's own force, grey propaganda is intentionally unattributable, and black propaganda is intentionally attributed to false parties. When dealing with a recalcitrant enemy, the use of all three types of propaganda might be necessary.
9. For an in-depth analysis of civilian in-theatre interpretations of collateral damage, see Janina Dill (2019).
10. I would like to especially thank LTC Bob Underwood, MAJ Scott Orr, Seumas Miller, and Adam Henschke for their assistance in helping to develop some of the ideas in this chapter.

REFERENCES

Bapat, Navin A. (2006), 'State Bargaining with Transnational Terrorist Groups', *International Studies Quarterly* 50 (1), 213–29.

Daniels, Norman (1979), 'Wide Reflective Equilibrium and Theory Acceptance in Ethics', *Journal of Philosophy* 76 (5), 256–82.

Department of the Army, and Department of the Navy (2003), 'Joint Publication 3-53, Doctrine for Joint Psychological Operations', 5 September.

Dill, Janina (2019), 'Distinction, Necessity, and Proportionality: Afghan Civilians' Attitudes towards Wartime Harm', *Ethics & International Affairs*, 3 (3), 315–42.

Gross, Michael, and Tamar Meisels (2017), *Soft War: The Ethics of Unarmed Conflict*, Cambridge: Cambridge University Press.

Johnson, Douglas A., Alberto Mora, and Averell Schmidt (2016), 'The Strategic Costs of Torture: How Enhanced Interrogation Hurt America', *Foreign Affairs,* September–October, accessed at https://www.law.upenn.edu/live/files/5734-johnson-mora-schmidt-the-strategic-costs-of.

Lazar, Seth, and Helen Frowe (2018), *The Oxford Handbook of Ethics of War*, New York: Oxford University Press.

Ludvigsen, Jan Andre Lee (2018), 'The Portrayal of Drones in Terrorist Propaganda: A Discourse Analysis of Al Qaeda in the Arabian Peninsula's *Inspire*', *Dynamics of Asymmetric Conflict*, 11 (1), 26–49.

Scharre, Paul (2016), 'American Strategy and the Six Phases of Grief', *War on the Rocks*, accessed at https://warontherocks.com/2016/10/american-strategy-and-the-six-phases-of-grief/.

Summers, Harry (1982), *On Strategy: A Critical Analysis of the Vietnam War*, New York: Random House Publishing.

Underwood, Bob (2019), 'Can Soldiers Justify Killing Some as a Means to Influence the Decisions of Others?', essay within the framework of the *5th Annual Oxford Uehiro Prize in Practical Ethics, Graduate Category*, 12 March, available at http://www.bioethics.net/2019/03/oxford-uehiro-prize-in-practical-ethics-question-can-soldiers-justify-killing-some-as-a-means-to-influence-the-decisions-of-others/.

United Nations Office of Drugs and Crime (2012), 'The Use of the Internet for Terrorist Purposes', accessed at https://www.unodc.org/documents/terrorism/Publications/Use_of_Internet_for_Terrorist_Purposes/ebook_use_of_the_internet_for_terrorist_purposes.pdf.

12. From ‘need to share’ to ‘need to care’: information aggregation and the need to care about how surveillance technologies are used for counter-terrorism

Adam Henschke

1. INTRODUCTION

Imagine this scenario: Anne, a soldier on deployment at a military base in a foreign conflict, meets a friend, Barry, for their morning jog. She posts, before and after, selfies of herself and Barry, updates her ‘JoggerLogger’ social media account with details of her run, and then heads to the shower. In doing this, Anne has put counter-terrorism operations at risk. The underpinning problem is that Anne has not treated potentially important information with due care. This chapter argues that individuals need to be careful with their personal information and that of others, even if that information is publicly available and/or relatively innocuous. Ultimately, I suggest that we need to shift our attitude to personal information from ‘need to share’ to ‘need to care’.

To explain, let us start with the shower. While taking a shower in and of itself is hardly a cause for alarm, in this scenario, Anne is using a shower with a heating system that is linked to a ‘smart meter’. To reduce energy use on the base, smart meters are being linked to smart grids to identify and anticipate peaks and lulls in energy use (Zhou et al. 2016). Every time the shower is used, the smart meter collects and communicates that spike of use. Recognizing that energy spike – and information on its timing, intensity and duration – can lead an external observer to infer that someone is taking a shower. Unfortunately, this smart meter was sold by a vendor who not only retained the default passwords (Chapman and Uren 2018; Kan 2016; Pishva 2016), but also had such poor cybersecurity practices that anyone with minimal cyber skills could find those passwords (Chapman and Uren 2018); thus, attackers are able to hack into the smart meter’s communications to gather information on use, and to

analyse the information for patterns. The smart meter gives the hacker information that forms a picture of the patterns of the base's life. For those wishing to understand the movements within a military base, the times when people shower can provide useful data for when to plan attacks, and when those on the base might be preparing for their own offensive operations. The point is that this new technology creates new opportunities for innocuous data to be gathered and analysed in ways that reveal sensitive information.

Anne's second mistake was taking selfies with Barry and posting them to her social media account. While photos of the two of them are great for friends and families to see, unbeknown to Anne, Barry is often involved in recruiting foreign assets for counter-terrorism Intelligence work. Being a uniformed soldier, Anne obviously has no concern about being identified as part of the military, but Barry has to be more careful. The problem here comes from the decline in price, ease of access and increased power of artificial intelligence (AI) to power facial recognition technology (FRT). Consider that an enemy Intelligence operative sees Anne in her uniform and takes a series of photos of her. FRT is used to identify her face, and AI is then used to trawl social media for her face. Every time another person comes up in Anne's photos, their face is identified and flagged as an associate. Barry is now flagged as an associate of Anne's, and his actions are put under closer attention. This makes Barry's counter-terrorism job much harder and potentially puts any locals Barry is seeking to recruit to the counter-terrorism operation at risk. Since he started working in counter-terrorism, Barry has been careful about his personal information, but old social media posts are found from the military training he did fifteen years ago. There are a number of photos of Barry and his friend, Claire, who trained together. Through FRT, Claire's face comes up in old photos. Unfortunately, Claire is currently running a secret counter-terrorism operation under a fake identity. This operation and Claire's life are now at risk. The new technologies pose a real risk to effectiveness, operational security and individual safety.

Finally, there is Anne's JoggerLogger. This is an imaginary brand of a wearable technology that monitors Anne's heart rate and other personal vitals, as well as locating her jogging times and routes. Being a social networking company, JoggerLogger posts all this information to the JoggerLogger community for them to compare and motivate each other to be their best. Problematically, this particular technology and company are big supporters of national security, and give a 50% discount to active military, police and other national security employees, meaning that it is the favoured device and platform of members of these communities. As such, a canny observer can guess where groups of military, police and national security people are located by identifying clusters of joggers on the JoggerLogger global map. The point here

is that certain technologies and their integration with social media can become uniquely sensitive if a particular pattern of use or user is identified.

None of this should be at all surprising. That we can induct something like shower use from other information, such as spikes in energy use, is hardly a shock. Likewise, the concerns of security officials about social media and its impacts on undercover operations have been publicly discussed since at least 2015 (Lord 2015). The JoggerLogger example is slightly adapted from a case in 2017. In this case, a wearable device associated with jogging was connected to the Internet, uploading the data to a publicly accessible website, Strava. The fitness-tracking app revealed potentially sensitive information about military bases and supply routes via its global heat map website. The data map shows one billion activities and three trillion points of latitude and longitude from 'Strava's global network of athletes…according to the American company… Using satellite imagery, you can see base buildings, for example. But on the heat map, you can see which buildings are most used, or the jogging routes of soldiers' (Bogle 2018). This security weakness was not particularly complex and was exposed by a master's student.

What is new and relevant here are the technologies, what they can reveal through the aggregation of seemingly innocuous information, and the pressure they put on how we understand and treat personal information. Because these technologies lead to a capacity for aggregation of innocuous information, this creates problems morally and for counter-terrorism. In this chapter, I promote the general idea that information – particularly innocuous information – should be treated with care. I offer the conceptual mechanics that underpin this claim.

2. SHIFTING RELATIONS TO INFORMATION: FROM 'NEED TO KNOW' TO 'NEED TO SHARE'

Until late 2001, in the US and elsewhere, national security agencies followed a general rule in the way they treated sensitive information: one only got access to it on a 'need-to-know' basis. As a US Congressional Research Service report puts it: 'The basic approach taken by the U.S. Government has been focused on establishing "need-to-know." Sensitive information is made available only to those persons with appropriate clearances and a "need-to-know" that information for the performance of their duties' (Best 2011).

Then, on 11 September 2001, the US suffered its worst domestic terrorist attack, and their national security infrastructure changed. As came to light, many of those who hijacked the planes were on various watch lists (National Commission on Terrorist Attacks Upon the United States 2004, pp. 83–4). The question then became, if the state knew that these people posed a threat, how did they slip through the net? One of the key conclusions drawn from the 9/11 investigation was that, though some arms of the US national security

apparatus knew about these potential threats, this information was not shared with its other arms. A key weakness was identified – that information relevant to national security was not being effectively shared across the vast body of national security agencies in the US: 'In the aftermath of the 9/11 attacks in 2001, a consensus emerged that information sharing, especially between Intelligence offices and law enforcement officials had been deficient and had contributed to the failure to detect the plot in advance' (Best 2011).

'Need to know' had prevented internal sharing of information. Because of the 9/11 attacks, there was a deliberate internal shift in the ways that sensitive information was to be treated. 'Need to know' was no longer the default. The US shifted its position from 'need to know' to 'need to share' (Best 2011). Responding directly to Intelligence failures brought about by information restriction, the default position became more active sharing of information. In parallel, by 2017, more than four million people in the US were eligible to access confidential, secret or top-secret information (Office of the Director of National Intelligence 2017). After the 9/11 attacks, the national security community saw a 'need to share' more information more freely to prevent another such attack from occurring. According to Genevieve Lester (2016), this sort of shift is common in Intelligence practices – there is a pendulum that swings between increased oversight and constraint and greater scope for freedom and power following tragedies (pp. 162–63). Following the 9/11 attacks, more information was being shared by more people more easily.

Parallel to these changes in attitude in the Intelligence communities, we have seen a similar attitudinal shift in the public at large. Many of us now actively and willingly share vast amounts of personal information on social media:

> What marks this age as one of *surveillance* is our own role in this – it is not simply that there are new information technologies…we are often the willing sources of this information, happily uploading selfies, buying wearable surveillance technologies, actively publicising [p]ersonal [i]nformation like no other time in history. (Henschke 2017, p. 4, emphasis in original)

Moreover, those social media and information companies have led to the development of so-called 'surveillance capitalism', where private companies make billions of dollars through the information that we provide to them (Zuboff 2019). We now place so much personal information into the public realm that the information once collected by police states seems quaint. Moreover, any claims to privacy seem confused if we are the active sources of the information (Henschke 2017). In short, individuals' behaviour and the modern economy are all evidence of a widespread attitude that we 'need to share' our personal information.

Given these institutional and social shifts towards massive sharing of personal information, often in public spaces, what does this mean for practices like counter-terrorism surveillance? One inference made by some is that privacy is dead – there is so much personal information 'out there' that we need no longer worry about adhering to privacy. Another upshot is that those working in national security sectors like counter-terrorism need to take better care with their own personal information. However, as we look at different notions of privacy, we will see that the first implication is conceptually muddled. Moreover, as we look at the revelational powers of these new technologies, we will see that, not only do those working in areas like counter-terrorism need to take better care with their own information, but they also need to take more care with other people's information.

3. RETHINKING PRIVACY[1]

The technological challenge to notions of privacy is central to the discussion and requires us to engage with the tight relation between privacy and technology. The liberal-democratic concept of privacy was crystallized in the seminal paper 'The Right to Privacy', written by Samuel Warren and Louis Brandeis in 1890 (Warren and Brandeis 1890). Importantly, this concept was developed *in response* to new technologies: 'In the late 19th century cameras had become portable, could take photographs practically in an instant and could be used by almost anyone who could afford one. Foreshadowing current debates about surveillance technologies, Warren and Brandeis were concerned about the ways that new technologies invaded personal space' (Henschke 2017, p. 35). This 'new [photographic] technology made it important to explicitly and separately recognize this protection under the name of privacy' (DeCew 2006). In the liberal-democratic tradition, at least, technology and privacy have had a close relationship with modern notions of privacy being developed *in response* to new technologies. The point is that we should not assume that new technologies necessarily mean the death of privacy.

To make sense of this claim that privacy is still very much alive, we need to understand what privacy refers to. A common way to think of privacy is as something secret. This notion of privacy-as-secrecy takes its roots in ancient Greek thought, where a binary distinction was made between political and domestic life, the *polis* and *oikos* (Arendt 1958, p. 24). This binary, where privacy is understood in contrast to the public, leads to what Daniel Solove calls 'the secrecy paradigm': 'Under this view, privacy is violated by the public disclosure of previously concealed information' (Solove 2008, p. 21). Importantly, when privacy is understood as secrecy, 'when others know the information, it is no longer completely secret' (Solove 2008, p. 139). Thus, if a person willingly places personal information into the public sphere, it seems

strange for them to claim that people ought to respect their privacy. Likewise, once something is publicly accessible, it is no longer private, and so – on a simplistic application of the secrecy paradigm – that information is no longer afforded the protections of privacy.

However, privacy is more than simply secrecy. When thinking of it in a political sense, privacy is seen as the opposite to government intrusion: the private describes that zone that the government is not permitted to interfere in (Henschke 2020). Continuing this political frame, privacy might be thought of as an instrumental good, something necessary for democratic freedom (Greenwald 2014, p. 177). Taking it from the explicitly political, we might instead think of privacy as a space of non-interference. Privacy 'is a set of boundaries we create between ourselves and others' (Solove 2008, p. 74). We can also think of privacy as control, specifically, 'the control we have over information about ourselves' (Fried 1969, p. 482). Here, privacy draws from the recognition that an individual has some legitimate claim to control their personal information. Another view suggests that, while 'control' is morally important, privacy is better understood as being concerned with access (Macnish 2018).

More recent accounts take pluralistic approaches, arguing that we think of privacy in different terms, such as data protection (van den Hoven 1999), or 'context-relative informational norms' (CRINs) (Nissenbaum 2009), or that privacy is a bundle of related concepts (Henschke 2017, pp. 28–55). The data protection account seeks to avoid unnecessary conceptual debates about what privacy *is*, and instead focuses on the *ends* of privacy: it asks what privacy is actually doing for us and why access to information should be constrained (van den Hoven 2007, p. 320) by identifying four moral justifications for protecting data: '1) Information-based harm; 2) Informational inequality; 3) Informational injustice; and 4) Encroachment on moral autonomy' (van den Hoven 2007, p. 320). In a similar line of reasoning, Helen Nissenbaum argues that we should respond to privacy concerns not by reference to some particular conception of privacy, but instead we should be concerned with determining appropriate information flows. 'Usually, when we mind that information about us is shared, we mind not simply that it is being shared but that it is shared in the wrong ways and with inappropriate others' (Nissenbaum 2009, p. 142). She looks at CRINs: these are 'characterized by four key parameters: contexts, actors, attributes and transmission principles' (Nissenbaum 2009, p. 140). In other writing, I have suggested that we need to see both descriptive and normative concepts play a role in a broader pluralistic idea conception of privacy (Henschke 2017, pp. 28–55).

This list is not exhaustive.[2] It does not claim to capture all the myriad concepts of privacy and their interactions.[3] Moreover, it does not aim to resolve which of these concepts is the correct one – quite the contrary. Part of the

problem with our current understanding of privacy is a search for the correct concept. Consider the opening paragraph from Julie Inness's (1992) *Privacy, Intimacy, and Isolation*:

> Exploring the concept of privacy resembles exploring an unknown swamp. We start on firm ground, noting the common usage of 'privacy' in everyday conversation and legal argument. We find intense disagreement about both trivial and crucial issues... we find chaos...the ground starts to soften as we discover the confusion underlying our privacy intuitions. (p. 3)

My point here is twofold. First, we need to recognize that there are a range of ways that we can understand privacy, and these extend far beyond seeing privacy simply as secrecy. Thus, we have a range of conceptual tools at our disposal to understand and apply to the production, collection and use of personal information. Second, just as national security communities changed their attitudes to information following the 2001 attacks, as technologies and our behaviours continue to evolve, we need to change attitudes to personal information again.

One way to start this attitudinal shift is to think of personal information as being concerned not just with what is in public or private, or even who controls or has access to the given information, but whether that information is intimate, or from a terrorism/counter-terrorism perspective, sensitive. Under a conception where privacy is concerned with intimacy, the starting point is the relation that an individual has to certain personal information. Specifically, an intimacy account holds that what is of relevance is a person's attitudinal stance – that they like, love or care about particular information:

> When an agent characterizes an act or activity as intimate, she is claiming that it draws its meaning and value from her love, liking or care. Intimate decisions concern such matters and, thus, involve a choice on the agent's part about how to (or not to) embody her love, liking or care. (Inness 1992, pp. 74–5)

On Inness's account, privacy is an attitudinal state whereby those decisions, actions or facts about a person which they love, like or care about are what is of interest.

I suggest here that national security communities take a similar approach to information – they recognize that certain information is *sensitive* and ought to be treated in a particular way because of that sensitivity. The basic idea of sensitivity is that, due to the importance of information for reasons such as national security, Intelligence or that it is relevant to an ongoing counter-terrorism operation and so on, those tasked with using or controlling access to that information now have a particular attitudinal stance towards it. Information deemed sensitive in a national security context is often classified

as confidential, secret, top secret and so on. As a result of these classifications, those working with it treat that information with due care, and have a set of processes in place to ensure that it continues to be treated with due care.

The public/private distinction and notions of secrecy are not of primary concern here; what is of importance is our attitude to that information, and how that attitude shapes our access to, and use of, that information. This notion of caring for information, showing the proper attitude towards information that recognizes it might be intimate or sensitive, is not just relevant in a general moral sense but for counter-terrorism practices as well (see below). However, we need to make one more step before we can see why personal information, particularly seemingly innocuous personal information, needs to be treated with care.

4. ANALYTICS AND REVELATION

The claim that we ought to treat certain information as intimate (when in a personal context) or sensitive (when in a national security context) with due care may be obvious. However, given the power of information technologies to collect and analyse vast amounts of information to produce new and increasingly intimate and sensitive information, we need to treat seemingly innocuous information with care. Consider that a teenage girl buys the following items: cocoa-butter lotion, a large purse, vitamin supplements (zinc and magnesium) and a bright blue rug. Now imagine that the girl's family subsequently receives a package in the mail congratulating her on becoming pregnant. The company, Target, did this. They had been using data analytics to reveal useful information about their customers, such as their 'pregnancy score' (Hill 2012). The point here is that what seems like mundane information when analysed can be particularly revealing. It can expose or uncover things about a person that are particularly intimate or sensitive, despite the initial information being innocuous or mundane.

This is the key observation from the opening example about Anne posting selfies, using wearable technology that communicates her actions with social media and using a smart-metered shower: each of these actions and the information they produce alone are innocuous and mundane. However, when particular technologies are applied to those actions, sensitive information can be produced or revealed. As I have argued elsewhere, the aggregation and analysis of innocuous information can reveal intimate and sensitive information, and can create new information from that mundane information that is highly revealing (Henschke 2017, pp. 144–49). The point is that, due to the revelational power of these new technologies, we need to treat even innocuous and mundane information – given it is aggregated and analysed – with increased care.

The concern is that in assessing data points independently of each other, we make a 'mistake in our moral mathematics' (Parfit 1987, pp. 67–86). The moral importance of a particular action is undervalued as a result of considering it independently:

> It is not enough to ask, 'Will my act harm other people?' Even if the answer is No, my act may still be wrong because of its effects. The effects that it will have when it is considered on its own may not be its only relevant effects. I should ask, 'Will my act be one of a set of acts that will *together* harm other people?' The answer may be Yes. And the harm to others may be great. If this is so, I may be acting very wrongly. (Parfit 1987, p. 86, emphasis in original)

Given the increased ubiquity of information technologies, and their increased capacities to analyse and reveal sensitive information, what we need to ask is *whether the sets of data together* will harm other people. Purchasing cocoa butter is of almost no consequence. Being pregnant is not. Taking a shower is largely irrelevant. The behavioural patterns that it generates can reveal militarily sensitive information, which is highly important in a conflict zone. This is the core recognition of the shift from 'need to know' to 'need to share': we gain new information by the aggregation of existing information, and our attitudes also need to shift.

The power of sharing information comes from the ways in which information analytics lead to revelation. Through aggregation and analysis, new information is revealed and produced (Henschke 2017, pp. 126–51). Like the difference between a jigsaw puzzle before and after completion, aggregation and analysis afford a whole portrait to emerge. The power of analytics comes from converting the innocuous to the intimate, revelation of the profound from the mundane. What was largely irrelevant, in combination and following analysis, can become highly sensitive.

Combining this capacity for revelation with the conceptualization of privacy-as-secrecy, we ought to now be able to recognize the core point of this chapter: individual data points are innocuous, and their location in the public realm means that they are no longer secret. So why should we care about them? First, they can be easily aggregated and analysed to reveal intimate and sensitive information. Second, that because this information is sensitive – that is, could be detrimental to national security and so on – if it gets into the wrong hands, means that it ought to be cared about. As we saw, privacy is more than secrecy, so whether that sensitive information is in the public sphere is irrelevant. What is relevant is what it reveals, and we ought to treat it as important. Maybe one could claim that we always need to have a duty of care in relation to confidential information, however, the problem is that the potential for aggregation and analysis means that there is a need for a duty of care in relation to information in the public domain because it can be aggregated and analysed

in ways that enable harm. In short, we need to shift from 'need to share' to 'need to care'.

5. THE 'NEED TO CARE' FOR INFORMATION AND ITS IMPLICATIONS FOR COUNTER-TERRORISM

As with the shift in attitude from 'need to know' to 'need to share', I am suggesting that we should now make further changes in our treatment of personal information. Seeing privacy beyond the secrecy paradigm encourages an attitudinal shift to the way personal information is treated. Our attitudes should shift from 'need to share' to 'need to care'. We now have a theoretical apparatus to explain why we need to treat information with care: innocuous information is potentially revelational if aggregated and analysed. Insofar as what is revealed may be intimate or sensitive is something of moral and practical importance, it follows that we need to treat information with due care. In short, we can see that information, even if it is accessible and thus not secret, can and should be considered private and so ought to be treated carefully.

There are four general implications of this shift to 'need to care'. As said, the point is that we need to change our attitudes towards information, recognizing that innocuous information can be intimate and sensitive. For individuals, the first implication is that we take care with how information about *us* is collected, produced and used. Such a demand applies to what *we* post online, and what we allow companies and even governments to do with that information. Insofar as we are concerned about others treating information with care, we ought to be careful with information about ourselves that we make public.

The second implication for individuals arises from basic consistency – if we generally do not want others to access and use intimate information about us, then we ought not access and use intimate information about them. That is, we 'need to care' for their information, even if it is in public. Privacy, on a complex pluralistic notion, holds us to consider that innocuous information, even if it is about other people, is still due respect. Again, our attitudes need to shift such that even shared information is treated with care.

The third implication applies to those in the national security and counter-terrorism space. Because innocuous information can reveal sensitive information, those in the counter-terrorism space need to be careful with their own information. The opening example about Anne took a range of technologies to show how standard public sharing behaviours can pose national security risks and can undermine counter-terrorism efforts. The point, again, is that those involved in these areas and operations need to take special care with information. Normal sharing behaviours, such as taking and posting selfies, using wearable exercise devices and so on, need to be revisited when

in a context like counter-terrorism. This responsibility to care for information also applies to issues like procurement – one of the security vulnerabilities identified stemmed from the lack of effective security on smart meters for showers. The responsibility here is for those involved in things like logistics, procurement and so on to be particularly careful about the security vulnerabilities that can arise from innocuous information.

The final point is that, just as individuals need to take more care with what sorts of public information they access, so too do national security and counter-terrorism operations need to treat publicly available information with due care. The point is not to say that counter-terrorism operations that engage in surveillance are unjustified – given certain national security threats, privacy can be overridden. Rather, the point of shifting to 'need to care' is to show that justifications are still needed even when accessing publicly available information or innocuous information. State surveillance programs, even those that use publicly accessible information, require justification and independent oversight. Warranting processes, for instance, might be a way of ensuring that this information is treated with due care. As a guiding principle, the 'need to care' rule makes those working with personal information, particularly those working with innocuous personal information and/or publicly available information, see that such information still deserves to be seen as private.

6. CONCLUSION

To conclude, public information might still be considered under the umbrella of privacy, and innocuous information can be highly revealing. These points are vital to recognize as the revelational power of analytics, coupled with the ubiquity of surveillance technologies and pervasiveness of publicized behaviours, means that we are drowning in innocuous information. Yet, despite this information not being secret, we need to take care with how we treat it. Recognizing the plurality of privacy concepts allows us to think beyond the secrecy paradigm; seeing that personal information can be intimate or sensitive signals to those involved in national security that the information needs to be treated with due care. In short, we need to shift attitudes from 'need to share' to 'need to care'.

NOTES

1. This section draws from 'On Privacy', in *Ethics in an Age of Surveillance* (Henschke 2017, pp. 28–55).
2. Judith DeCew's (2006) privacy entry in the *Stanford Encyclopedia of Philosophy*, Daniel Solove (2008) and Helen Nissenbaum (2009) all give great overviews of the range of privacy conceptions.
3. See Koops et al. (2016) for more on this.

REFERENCES

Arendt, Hannah (1958), *The Human Condition, Charles R. Walgreen Foundation Lectures*, Chicago: University Of Chicago Press.

Best Jr., Richard A. (2011), *Intelligence Information: Need-to-Know vs. Need-to-Share*, Washington, DC: Congressional Research Service.

Bogle, Ariel (2018), 'Strava Has Published Details about Secret Military Bases, and an Australian Was the First to Know', ABC News, 30 January 2018, http://www.abc.net.au/news/science/2018-01-29/strava-heat-map-shows-military-bases-and-supply-routes/9369490.

Chapman, Eliza, and Tom Uren (2018), *The Internet of Insecure Things*, Canberra: Australian Strategic Policy Institute.

DeCew, Judith (2006), 'Privacy', *Stanford Encyclopedia of Philosophy*, accessed 19 September 2007 at http://plato.stanford.edu/archives/fall2006/entries/privacy/.

Fried, Charles (1969), 'Privacy', *Yale Law Journal* 77 (3), 475–93.

Greenwald, Glenn (2014), *No Place to Hide: Edward Snowden, the NSA, and the U.S. Surveillance State*, New York: Metropolitan Books.

Henschke, Adam (2017), *Ethics in an Age of Surveillance: Virtual Identities and Personal Information*, New York: Cambridge University Press.

Henschke, Adam (2020), 'Privacy, the Internet of Things and State Surveillance – Handling Personal Information within an Inhuman System', *Moral Philosophy and Politics* 7 (1), 123–49.

Hill, Kashmir (2012), 'How Target Figured Out a Teen Girl Was Pregnant before Her Father Did', *Forbes*, 16 February 2012, accessed 11 January 2016 at http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/.

Inness, Julie C. (1992), *Privacy, Intimacy, and Isolation*, New York: Oxford University Press.

Kan, Michael (2016), 'IoT Botnet Highlights the Dangers of Default Passwords', *InfoWorld*, 4 October 2016.

Koops, Bert-Jaap, Bryce Clayton Newell, Tjerk Timan, Ivan Skorvanek, Tomislav Chokrevski, and Masa Galic (2016), 'A Typology of Privacy', *University of Pennsylvania Journal of International Law* 38(2), 483–575.

Lester, Genevieve (2016), *When Should State Secrets Stay Secret?*, Cambridge: Cambridge University Press.

Lord, Jonathan (2015), 'Undercover Under Threat: Cover Identity, Clandestine Activity, and Covert Action in the Digital Age', *International Journal of Intelligence and CounterIntelligence* 28 (4), 666–91, doi: 10.1080/08850607.2015.1022464.

Macnish, Kevin (2018), 'Government Surveillance and Why Defining Privacy Matters in a Post-Snowden World', *Journal of Applied Philosophy* 35 (2), 417–32.

National Commission on Terrorist Attacks Upon the United States (2004), *Final Report of the National Commission on Terrorist Attacks Upon the United States*, Washington, DC: Government Printing Office.

Nissenbaum, Helen (2009), *Privacy in Context: Technology, Policy, and the Integrity of Social Life*, Stanford, CA: Stanford Law Books.

Office of the Director of National Intelligence (2017), *Fiscal Year 2017 Annual Report on Security Clearance Determinations*, Washington, DC: National Counterintelligence and Security Center.

Parfit, Derek (1987), *Reasons and Persons*, Oxford: Oxford University Press.

Pishva, Davar (2016), 'Internet of Things: Security and Privacy Issues and Possible Solution', *ICACT Transactions on Advanced Communications Technology* 5 (2), 797–808.

Solove, Daniel (2008), *Understanding Privacy*, Cambridge, MA: Harvard University Press.

van den Hoven, Jeroen (1999), 'Privacy and the Varieties of Informational Wrongdoing', *Australian Journal of Professional and Applied Ethics* 1 (1), 30–43.

van den Hoven, Jeroen (2007), 'Privacy and the Varieties of Informational Wrongdoing', in John Weckert (ed.), *Computer Ethics*, Aldershot: Ashgate Publishing, pp. 317–30.

Warren, Samuel D., and Louis D. Brandeis (1890), 'The Right to Privacy', *Harvard Law Review* 4 (5), 193–220.

Zhou, Bin, Wentao Li, Ka Wing Chan, Yijia Cao, Yonghong Kuang, Xi Liu, and Xiong Wang (2016), 'Smart Home Energy Management Systems: Concept, Configurations, and Scheduling Strategies', *Renewable and Sustainable Energy Reviews* 61, 30–40, doi: 10.1016/j.rser.2016.03.047.

Zuboff, Shoshana (2019), *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, London: Public Affairs.

13. Bulk data collection, national security and ethics

Scott Robbins

1. INTRODUCTION

The rise of Internet communications has necessitated a rise in digital national security intelligence collection (including counter-terrorism intelligence and military intelligence) – currently at a scale never seen before in liberal democracies. The Snowden revelations of 2013 exposed digital intelligence collection that was pervasive and perhaps illegal (Greenwald 2013). People around the world were shocked at the capabilities of the National Security Agency (NSA), and the intelligence-collection practices revealed by Snowden have not slowed. On the contrary, many of these practices are being enshrined in the law (Pieters 2016; Travis 2016). Whether or not these practices are legal, it is essential to understand whether or not they are ethical – or how these practices can be conducted ethically. This involves identifying what makes these practices different from those that came before. Then, one must highlight how this changes the ethical analysis.

Two broad ethical paradigms constrain the practice of intelligence. First, there is what is acceptable for law enforcement – which generally takes a case-by-case approach to evaluating the acceptability of collecting intelligence or surveilling a subject or subjects. Essential considerations for law enforcement are: that there is reasonable suspicion or probable cause that the suspect (or suspects) are going to commit a serious crime, that the intrusion of their privacy is not disproportionate to the violations of citizens who will be the victims of that crime, and that the intelligence collection is necessary (that is, there is no less-intrusive alternative). Second, there is what is acceptable for national intelligence agencies to do during war, a context in which there are few constraints on their intelligence collection and analysis activities. Bulk data collection (BDC) for counter-terrorism purposes poses problems for each of these paradigms for two reasons. First, since terrorism is a crime, it can and should be dealt with as a crime by law enforcement (Miller 2008); however, terrorist groups, such as the Islamic State (ISIS), have at times launched

military-style campaigns that target the state as a whole and, therefore, may require wartime tactics in response. Second, by definition, BDC sweeps up large amounts of data on innocent people, which is not something typically allowed by law enforcement. This has created a murky situation concerning counter-terrorism intelligence collection and analysis. This chapter cannot solve this complex problem or, rather, set of problems; however, it does provide some clarity on, and justification for, constraints that ought to be imposed on one specific form of intelligence collection: BDC.

Contemporary scholars have frequently discussed the ethics of intelligence activities within a Just War Theory (JWT) framework – those principles deemed necessary for the ethical initiation, conduct, and termination of war. Scholars have made efforts to modify JWT into a Just Intelligence Theory (JIT) (Bellaby 2012, 2016; Gendron 2005; Macnish 2014; Omand and Phythian 2013; Quinlan 2007). My focus in this chapter is to apply some of the latest work in JIT to BDC. Thus far, there has been no comprehensive ethical review of the practice of BDC for intelligence purposes.[1]

2. BULK DATA COLLECTION

To collect in bulk roughly means that the scope of collection will likely pick up many records that are not associated with current targets (Anderson 2016; National Research Council 2015). For example, the intelligence community (IC) may want all records associated with the current ISIS leader Abu Ibrahim al-Hashimi al-Qurashi. If the IC were only to collect records associated with him, then the IC is not collecting in bulk; instead, they are conducting a targeted collection. However, if the IC wants, for instance, all records coming into and out of Syria because they think many terrorists are operating there, then the IC is collecting in bulk. There are many Syrians whose data will be collected who are not engaged in terrorist acts and who do not even interact with terrorists. This is significant from an ethical standpoint because the IC is knowingly collecting data on innocent people and doing so on a large scale. This will be important for evaluating BDC in terms of the ethical principle of proportionality (see below).

BDC is done in two ways:

1. Bulk Interception: the practice of intercepting Internet communications data that are in transit.
2. Bulk Acquisition: the practice of acquiring bulk data from telecommunications and Internet companies.[2]

Bulk interception is accomplished by placing fibre optic splitters on telecommunications entry points. These fibre optic splitters copy the data and pass

it along to intelligence agency infrastructure. These data will be filtered – to ensure that the data collection meets legal requirements[3] and that as little irrelevant data end up on agency servers as possible.

Bulk acquisition works in two ways. First, intelligence agencies can simply ask (or force) third-party institutions to turn over data in bulk (that is, data resulting from the application of some filter). Second, intelligence agencies may have back-door access to third-party institution servers. The Snowden revelations revealed that such back-door access was given to the NSA by Google, Facebook, and others (Greenwald and MacAskill 2013). In this chapter, BDC is also taken to be prima facie wrong, given that it involves infringing the privacy rights of innocent citizens on a large scale. The purpose of this chapter is to understand what conditions would have to be met for its use to be justified.

Privacy or other rights of any given *targeted* individual – that is, person who is an object of prior reasonable suspicion – cannot be the sole focus in the ethical evaluation of BDC. By definition, BDC is not targeted in this sense. Instead, it is the members of an entire group of people whose data will be collected to isolate members of the group for scrutiny. These groups are the result of filters being applied to the data passing through the Internet. The filters themselves, then, are where the focus should lie for an ethical evaluation of BDC. It is these filters that delimit the set of potential 'targets'. Few would have objected to a filter that selects all data related to Osama bin Laden. In the case of bulk collection, the filters are, by definition, much broader. These filters are what should be evaluated – in other words, this chapter focuses on understanding what might make the use of a particular filter morally justified and what might not.

3. JUST INTELLIGENCE

As already mentioned, a prominent theoretical perspective in the field of intelligence ethics advocates adapting JWT to evaluate intelligence collection and analysis. The primary reason for basing an ethics of intelligence on the ethics of war is that the conduct of both war and intelligence collection involves actions that are prima facie unethical. In war, you are killing people, destroying bridges and cities, holding people captive, and so on. All of these things are ethically bad. However, there are cases when such actions are necessary, proportionate, and morally justified. A country being invaded by another country should be able to defend itself – including shooting at their invaders. JWT outlines principles that are held to be necessary and sufficient to justify going to war (*jus ad bellum*) and to justify the conduct of that war once it is waged (*jus in bello*). Michael Quinlan (2007) argues that the practice of intelligence must also be justified and limited. In other words, there should be conditions that

justify starting an intelligence program and limitations on how to conduct that intelligence program justly. Quinlan (2007) names these *jus ad intelligentiam* and *jus in intelligentia*.

The reason for using a theory based on JWT for intelligence collection is that intelligence collection involves harm and/or rights infringements that need further justification. Intelligence collection can involve listening in on private conversations, torture, deception, interception of communications, and so on. All of these actions would also be ethically disallowed under normal circumstances.

Harms from BDC can be divided into two types: privacy infringements and restrictions on autonomy. The data swept up by an intelligence agency belong to someone. An individual owns the information that those data reveal (Bellaby 2012). Prima facie, no one should be allowed to take these data. Of course, if this person is a known terrorist, then a state would be justified in collecting all information about this person. The point is that a state needs to justify its actions concerning BDC because harm or rights infringements are associated with such intelligence programs. If the state fails to justify such infringements, then violations have occurred.

People's autonomy – including citizens of the bulk-data-collecting state – can be restricted – intentionally or unintentionally – by BDC programs. Public knowledge of government BDC could affect innocent people's autonomy whether or not their data are collected. The so-called 'chilling effect' is when governmental regulation and policy not directed at certain activities deters individuals from carrying out protected activities (Robbins and Henschke 2017).

4. JUST BULK DATA COLLECTION

It is not the purpose of this chapter to justify the use of JIT; rather, it is to use principles of JIT to tease out ethical issues that arise due to the practice of BDC. In what follows, I use the JIT principles of just cause, proportionality, right intention, and proper authority to uncover issues that must be overcome to justify BDC's use.

4.1 Just Cause

What would be a just cause for intelligence collection? As counter-terrorism is the most salient reason given in recent times for BDC, this analysis will be restricted to cases involving terrorism.[4] At first glance, it is clear that counter-terrorism is a just cause for an intelligence operation. If terrorists are attempting to conduct attacks on citizens of a country, that country has just cause to collect intelligence that would prevent those attacks. Arguably, things

might not be so simple for the reason that 'the general threat of terrorism, the so-called War on Terror, for example, is too indistinct to offer any specific just cause for an operation' (Bellaby 2016, p. 313).

Someone might claim that the IC has just cause to collect intelligence on everyone in the world to prevent unknown future threats from being realized. Since the IC does not know where the threats could come from in the future, no restriction on data collection would occur. This argument is spurious even if one is working with a reasonably broad definition of national security.

However, this is not a complete picture for two reasons. First, BDC occurs on a spectrum. At one end of the spectrum are practices of BDC that are unquestionably targeted; at the other end are practices of BDC that cannot be in any way considered targeted. At the end of the spectrum where the most-targeted practices are conducted, there might be practices such as collecting all the available data about the citizens who live in a small town known to be the home of some terrorists. At the other end of the spectrum might be practices such as collecting all the available data about United States citizens and anyone else who has entered the United States or who has communicated with anyone who lives in the United States. The justification for a particular instance of BDC will depend in part upon where it falls on this spectrum. Second, there is a conceptual issue regarding the point at which intelligence has been collected. On one account (further explained below), it seems as if the NSA, for example, collects most of the data travelling through the Internet as most of the world's data is routed at some point through the United States – although there are attempts to change this.[5] On the NSA's account of collection, the NSA collects a tiny fraction of the data travelling through the Internet. The result of this analysis will affect when the just cause principle can be applied.

Now comes the conceptual issue of what counts as 'collection', as it is not simple in the case of BDC. When can data be said to have been 'collected' by an intelligence agency? It may be helpful to take a rudimentary look at an email that ends up in the hands of an intelligence analyst through BDC:

When the email is sent, it gets routed to the backbone of the Internet run by (mostly) US communications companies (for example, AT&T).[6] The communications company acts as the post office in that it makes sure the communication is directed towards the intended recipient. It is here, at this first stage of the process (Stage 1), that, for example, the NSA has a splitter on the fibre optic cables to copy the data. At this stage of the process, the data would have to be stored until filters could be run on it. At the next stage of the process (Stage 2), the filters go through the data, discarding information that does not match any of the filters. At Stage 3, the data that make it through the filters end up on NSA servers for storage. Finally, at Stage 4, an analyst queries the

data, resulting in the email (along with other data, perhaps) being returned to the analyst who reads it.

With Stage 1, above, it is clear that, for some time, the email is stored on a government server. NSA-owned equipment has possession of the data; however, at least as I have described the process, there is no potential for analysts to access those data.[7] An example from the physical world may help clarify the point. When you put your bag on the conveyor belt, it now sits on airport security property. If the machine that selects baggage for inspection were automated (with no human in control), this would be much like the BDC situation. All bags must pass through, but only a few are passed on for further inspection. We would hardly say that our bags have been collected (or that our privacy has been infringed) simply because they are on the conveyor belt. But once that bag is directed away from all the other bags towards the inspection team, the bag has been 'collected' (Stage 3). In this analogy, the bag going through the machine is like the data in temporary storage – it rests on the collector's property. Still, it is inaccessible to them (again, provided that the baggage machine is automated).

The intervention at Stage 2 appears trivial at first glance. It is merely the state of the data as filters are being run on them. It should look like a series of questions: Did this data come from Syria? No. Iraq? No. Is it encrypted using tools known to be used by terrorists? No. And so on. If any of the questions results in a yes, then the data move on to long-term or permanent storage. I include Stage 2 in my discussion because I want to highlight the difference between using these filters and running complex pattern-matching algorithms. Filters appear to be merely automating a human process. If one were to print out all of the emails passing through the Internet, a human could, in principle, check to see which of the emails matched one of the filters. Computers speed this process, but a human being could easily double-check each communication if need be. This is opposed to complex computer algorithms attempting to find patterns in the data or make predictions on the data. For example, a deep-learning algorithm could be trained on all of the communications associated with terrorism (previously) and used to classify future communications in terms of the connection with terrorist communications or other terrorist actions. This is no longer the automation of a human process; rather, it is a novel process that would be opaque to human minds. What can be said about algorithms like these being run on the data in temporary storage? Earlier I argued for evaluating the filter for just cause, but in this case, the filter is opaque to evaluation. The computer scientist who created the original algorithm would not even be able to explain how the algorithm classified a particular communication as being associated with terrorism. We would lack meaningful human control over how the algorithm selects communications to collect (Robbins 2019a; Santoni de Sio and van den Hoven 2018).

While the filter cannot be evaluated in the case of a machine-learning algorithm, an argument could be made that, if the algorithm is better at classifying communications in terms of their connection to terrorism than the articulable filters are, then the fact that they are not articulable should not be a reason not to use them. In other words, using machine-learning algorithms could be better for privacy because they are more accurate in their classifications. A similar point has been made in other contexts (Esteva et al. 2017; Robbins 2019b). This argument, however, fails in the context of counter-terrorism. First, the reason that the IC is collecting data in bulk is in part because of the changing communication tactics of terrorist groups. The classification of communications into those relevant to terrorist activity, and those not relevant, will change drastically over time. This is so for three reasons: first, as technology changes, the way we, as a society, communicate changes; second, terrorist groups of the future may communicate drastically differently than terrorist groups of the past; and third, terrorist groups know they are being surveilled and modify the way they communicate to thwart intelligence agencies. Therefore, it will be challenging to say that an algorithm is better at classifying communications than is an articulable filter.

At Stage 3, it is common to classify the data as collected. In this case, the data rest on government servers with access given to analysts under institutional constraints. These data are justifiably collected when there is evidence that there is a terrorist threat being organized or planned by the group described in the filter resulting in the collected data *and* that this threat is directed at the state collecting those data. While this may satisfy just cause, whether or not it is proportionate to the threat is another question.

4.2 Proportionality

Proportionality is a comparative notion where we judge that 'an act is wrong if the relevant harm it will cause is out of proportion to its relevant good' (Henschke 2018). Talk of proportionality with respect to going to war (*jus ad bellum*) is stated as a condition that 'the destructiveness of war must not be out of proportion to the relevant good the war will do' (Hurka 2005, p. 35). The principle of proportionality is also used for evaluating the just conduct of war (*jus in bello*), albeit in the context of the principle of discrimination and the principle of military necessity. According to the principle of necessity, the action must serve a military purpose. According to the *jus in bello* proportionality principle, the (unintended) deaths of innocent civilians, while permissible if militarily necessary, must not be disproportionate in the sense that the number of innocent deaths is disproportionate relative to the importance of the military objective (Hurka 2005).

With BDC, one can quickly see that the evaluation of proportionality hinges on empirical data. For any proportionality calculation, 'we need specific facts about the costs [and] we need specification about the ends [that] are being sought' (Henschke 2018). The extent of the harm done by BDC is challenging to determine before it has been carried out. How pervasive is the chilling effect mentioned in Section 3 above? A Pew Research Center poll concluded that 25 per cent of Americans had changed their online behaviour due to perceived government surveillance (Gao 2015). Depending on the methods used, the harms could be even more widespread – and more difficult to quantify.

The bulk acquisition of data from third-party institutions – especially when it pertains to back-door access and data retention – could result in diminished trust in participating institutions. Edward Snowden, in an interview with *The New Yorker*, explicitly told people not to use Dropbox, Google, or Facebook because of their susceptibility to intelligence collection (*The New Yorker* 2014). This, in turn, could harm the profits of third-party institutions and the US economy itself.

It will be necessary going forward to understand the harms to third-party institutions as a result of BDC. Harms like these must be taken into account in any calculation of proportionality. These harms would then have to be weighed against the efficacy of the program – or the good that it will do, which of course, is another empirical matter.

4.3 Right Intention

If the government intends to prevent terrorism, then right intention should be of little concern. Much like just cause above, the situation is not so simple. There may be a clear threat in Afghanistan of terrorism directed at the United States – a threat that constitutes just cause for BDC. However, the intention of the collecting state may be to glean information helpful to influence elections there. If that were the case, then the collecting state does not meet right intention.

What complicates right intention, however, is when and how often it should be applied. Right intention should be applied to the decision to create a filter that results in BDC. However, there is a time dimension that complicates this in two ways: (1) the filter will continue to collect data long after the decision was made to use that filter, and (2) the collected data will be stored long after that decision.

To illustrate the problem with (1) above, let us act as if BDC was a tactic to combat the Irish Republican Army (IRA), and the British intelligence agencies had just cause to collect all of the data coming into and out of Ireland. The IRA is no longer the threat it once was, so not only would British intelligence have to re-evaluate just cause, but they may have a problem with right intention as

the British intelligence agencies may leave the filter because the data could be useful in the future. The time dimension of BDC means that the collected data should be tied to the justification for the creation of the filter – and deleted when that justification no longer holds.

4.4 Proper Authority

One could go along with traditional JWT and claim that BDC's only proper authority is the state. If this is the case, then a problem arises because, in practice, there are many third-party institutions collecting data in bulk. The practice of bulk acquisition is about the state copying data that have already been collected by third-party institutions – either by request or by back-door access. The question becomes whether or not the third party is then collecting bulk data as part of an intelligence collection and analysis program.

In many instances, this is not the case at all. Telecommunications and Internet companies store a lot of data that are necessary to conduct their business. Google does not store your email on their servers for reasons of national security; they store your email so that you have access to it. There is nothing inherently wrong with the IC obtaining data from third parties. If Osama bin Laden had a Gmail account, it would, and should, be expected that the NSA ask Google for those records – and it would, and should, be expected that Google provides them.

Things get more interesting if we look at forced data retention policies – in which laws mandate that third-party institutions retain data they may not typically retain for counter-terrorism (or national security). Now, the third-party institution is engaging in BDC as an intelligence program. This fails the principle of proper authority. Not only this, but now all of the data that have been retained that usually would not be should be included in our evaluations of just cause, right intention, and proportionality.

This problem is exacerbated when it is understood what the broad purpose of retaining such data would be. The purpose is, purportedly, national security. So the government faces a dilemma concerning the value of these data. Either the data are essential for national security, or they are not. If the data are essential, then the storage of those data should not be contracted out to third-party institutions. This is both because of the security risk of third parties being hacked and the blurring of institutional aims that such storage causes. Blurring these institutional responsibilities could damage the company's reputation and make it easier for those wishing to evade detection to choose other institutions. If the data are not essential, they should not force third-party institutions to retain such data.

5. CONCLUSION

This chapter has used JIT to evaluate the practice of BDC in liberal democracies for intelligence purposes. JIT forced the selection of an object of evaluation for BDC – in this case, the filters used to funnel data into government servers – and teased out some important ethical issues surrounding the practice. Most importantly, this evaluation pointed us to some essential constraints that should be placed on this practice. These constraints included: not using artificial intelligence as filters; the group specified by a particular filter must pose a threat to the collecting state; collected data must be tied to a filter and deleted when the justification for that filter no longer holds; and consumer companies, such as Google and Facebook, should not be allowed to act as intelligence agencies (collect data for the sole purpose of counter-terrorism).

This evaluation is just a start; however, it points to constraints that are not currently in place on BDC. Furthermore, this chapter starts from the premise that BDC is a valuable tool in the fight against terrorism. This may not be the case. If this tool turns out to be ineffective, it should not be used with or without the constraints outlined above. The point is that if intelligence agencies want this tool in their arsenal, they should be using it in a way that conforms to liberal-democratic principles and values. Having a just cause and right intention to collect data in bulk that are proportional to the threat and conducted by a proper authority would be a good start.

NOTES

1. Bellaby (2016) does give an in-depth ethical evaluation of cyber intelligence (broadly construed) with a couple of paragraphs on what he calls '*en masse* collection', the term he gives to BDC.
2. Bulk interception and bulk acquisition are terms used by David Anderson (2016) in his review of the UK's proposed Bulk Powers Act, which later became the Investigatory Powers Act.
3. In the United States, for example, there must be minimization procedures to ensure that as little US personal data as possible ends up on intelligence agency servers. See, for example, Blum (2008).
4. However, this analysis will apply to any context where national security is at stake.
5. See, for example, Edmundson et al. (2016).
6. This is not always the case, and increasingly there are methods for preventing your messages from being routed through surveillance states. See Edmundson et al. (2016).
7. Although if XKeyScore exists as described by *The Intercept* (Lee et al. 2015), then analysts *do* have access and the BDC program would fail to meet just cause.

REFERENCES

Anderson, David (2016), *Report of the Bulk Powers Review*, United Kingdom: Williams Lea Group, available at https://terrorismlegislationreviewer.independent.gov.uk/wp-content/uploads/2016/08/Bulk-Powers-Review-final-report.pdf.

Bellaby, Ross W. (2012), 'What's the Harm? The Ethics of Intelligence Collection', *Intelligence and National Security*, 27 (1), 93–111, https://doi.org/10.1080/02684527.2012.621600.

Bellaby, Ross W. (2016), 'Justifying Cyber-Intelligence?', *Journal of Military Ethics*, 15 (4), 299–319, https://doi.org/10.1080/15027570.2017.1284463.

Blum, Stephanie Cooper (2008), 'What Really Is at Stake with the FISA Amendments Act of 2008 and Ideas for Future Surveillance Reform', *Boston University Public Interest Law Journal*, 18, 269.

Edmundson, Anne, Roya Ensafi, Nick Feamster, and Jennifer Rexford (2016), 'Characterizing and Avoiding Routing Detours through Surveillance States', ArXiv: 1605.07685 [Cs], May, http://arxiv.org/abs/1605.07685.

Esteva, Andre, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, and Sebastian Thrun (2017), 'Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks', *Nature* 542 (7639), 115–18, https://doi.org/10.1038/nature21056.

Gao, George (2015), 'What Americans Think about NSA Surveillance, National Security and Privacy', Pew Research Center blog, 29 May, accessed 23 September 2019, at http://www.pewresearch.org/fact-tank/2015/05/29/what-americans-think-about-nsa-surveillance-national-security-and-privacy/.

Gendron, Angela (2005), 'Just War, Just Intelligence: An Ethical Framework for Foreign Espionage', *International Journal of Intelligence and CounterIntelligence*, 18 (3), 398–434, https://doi.org/10.1080/08850600590945399.

Greenwald, Glenn (2013), 'NSA Collecting Phone Records of Millions of Verizon Customers Daily', *The Guardian*, 6 June, accessed 23 May 2018, at https://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order.

Greenwald, Glenn, and Ewen MacAskill (2013), 'NSA Prism Program Taps in to User Data of Apple, Google and Others', *The Guardian*, 7 June, accessed 23 May 2018, at https://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data.

Henschke, Adam (2018), *Ethics in an Age of Surveillance: Personal Information and Virtual Identities*, Cambridge, UK: Cambridge University Press.

Hurka, Thomas (2005), 'Proportionality in the Morality of War', *Philosophy & Public Affairs* 33 (1), 34–66, https://doi.org/10.1111/j.1088-4963.2005.00024.x.

Lee, Micah, Glenn Greenwald, and Morgan Marquis-Boire (2015), 'A Look at the Inner Workings of NSA's XKEYSCORE', *The Intercept*, 2 July, accessed 5 April 2021, at https://theintercept.com/2015/07/02/look-under-hood-xkeyscore/.

Macnish, Kevin (2014), 'Just Surveillance? Towards a Normative Theory of Surveillance', *Surveillance & Society* 12 (1), 142–53.

Miller, Seumas (2008), *Terrorism and Counter-Terrorism: Ethics and Liberal Democracy*, Hoboken, NJ: Wiley.

National Research Council (2015), *Bulk Collection of Signals Intelligence: Technical Options*, Washington, DC: The National Academies Press, https://doi.org/10.17226/19414.

Omand, David, and Mark Phythian (2013), 'Ethics and Intelligence: A Debate', *International Journal of Intelligence and CounterIntelligence* 26 (1), 38–63. https://doi.org/10.1080/08850607.2012.705186.

Pieters, Janene (2016), 'Proposed Law Allows Massive Data Mining by Intelligence Agencies', *NL Times*, 15 April, accessed 15 May 2018, at https://nltimes.nl/2016/04/15/proposed-law-allows-massive-data-mining-intelligence-agencies.

Quinlan, Michael (2007), 'Just Intelligence: Prolegomena to an Ethical Theory', *Intelligence and National Security* 22 (1), 1–13, https://doi.org/10.1080/02684520701200715.

Robbins, Scott (2019a), 'AI and the Path to Envelopment: Knowledge as a First Step towards the Responsible Regulation and Use of AI-Powered Machines', *AI & SOCIETY*, 35 (2), 391–400, https://doi.org/10.1007/s00146-019-00891-1.

Robbins, Scott (2019b), 'A Misdirected Principle with a Catch: Explicability for AI', *Minds and Machines*, 29 (4), 495–514, https://doi.org/10.1007/s11023-019-09509-3.

Robbins, Scott, and Adam Henschke (2017), 'The Value of Transparency: Bulk Data and Authoritarianism', *Surveillance & Society* 15 (3/4), 582–89, https://doi.org/10.24908/ss.v15i3/4.6606.

Santoni de Sio, Filippo, and Jeroen van den Hoven (2018), 'Meaningful Human Control over Autonomous Systems: A Philosophical Account', *Frontiers in Robotics and AI* 5, Article 15, https://doi.org/10.3389/frobt.2018.00015.

The New Yorker (2014), 'The Virtual Interview: Edward Snowden – *The New Yorker* Festival', accessed 23 September 2019, at https://www.youtube.com/watch?v=fidq3jow8bc.

Travis, Alan (2016), '"Snooper's Charter" Bill Becomes Law, Extending UK State Surveillance', *The Guardian*, 29 November, accessed 3 May 2018, at http://www.theguardian.com/world/2016/nov/29/snoopers-charter-bill-becomes-law-extending-uk-state-surveillance.

14. Collective moral responsibility and chemical, biological, radiological and nuclear terrorism: the case of phosphine

Jonas Feltes

1. INTRODUCTION

The toxic gas phosphine has not been considered a pressing security threat by most security agencies in Western democracies.[1] This chapter will (1) determine the psychological and political impact of a terrorist attack with phosphine. Furthermore, it will be argued that (2) security agencies, as well as manufacturers and vendors of phosphine products, share a joint moral responsibility to deny terrorists access to this substance. After illustrating that (3) the current countermeasures in this regard are inadequate, this chapter (4) presents a prevention framework that enables all relevant actors to cooperate and, thereby, to fulfill their moral responsibilities.

2. PHOSPHINE AS A TERRORIST WEAPON

2.1 Properties and Toxicity

Phosphine is a colorless, toxic gas compound with the formula PH_3. It is heavier than air and exhibits LD_{50} values of 3.03 mg/kg.[2] Phosphine kills organisms by seriously disturbing the transport and use of oxygen in the body. Hence, it is considered a so-called pulmonary or choking agent (National Center for Biotechnology Information 2019). On an industrial scale, phosphine can be produced by acid-catalyzing white phosphorous, or by reacting white phosphorus with sodium or potassium hydroxide. However, for this chapter, the preparation of phosphine in small-scale applications is particularly interesting. For example, phosphine can be produced by mixing metal phosphides (such as aluminum phosphide or calcium phosphide) with water (Bogle et al.

2006; Gurjar et al. 2011; Gurusinghe 2014). This simple production of phosphine is deployed to use phosphine as a rodenticide.

2.2 Psychological and Political Damage

The psychological and political impact of a terrorist attack with phosphine has not yet been discussed in academia. However, there is academic discussion concerning chemical terrorism in general.

While the physical damage inflicted by a terrorist attack using chemical agents such as phosphine would likely be limited, a terrorist attack with an improvised chemical device would inflict widespread fear and panic among the affected population. This was particularly evident during the attack against the Tokyo subway in 1995 (Danzig et al. 2011, pp. 33–4; Parachini 2001, p. 391). The use of sarin in the underground infrastructure of Tokyo not only killed 12 people but also caused mass panic among subway passengers. Coupled with inadequate information about the nature of the attack, this anxiety caused over 5 000 people to seek medical attention due to actual or believed symptoms (Smithson and Levy 2000). The hospitals in the area quickly became overburdened with the number of new patients, of which the majority sought unnecessary medical help. Furthermore, the media coverage of the incident as well as decontamination efforts in the subway systems with personnel in hazardous materials suits added to this anxiety and caused people to avoid the area (for an argument that the media added to this anxiety in the aftermath of the attack, see Pangi 2002, p. 433).

Returning to our specific focus, it is essential to note that an attack with phosphine would not require any decontamination efforts above and beyond the ventilation of the affected area or building. However, as happened in Tokyo, it is likely that, at first, neither the victims of the attack nor the first responders will have any knowledge about the specific agent that was used in the attack. Hence, it is likely that fire departments and counter-terrorism forces will arrive on the crime scene with personal protective equipment (PPE) necessary for a chemical attack. The presence of responders with PPE would likely contribute to public anxiety. Media outlets and journalists would publish images of these decontamination operations and, thereby, increase the impact of the attack.

In addition to causing anxiety, a chemical attack in a Western democracy would cause political damage and damage to institutions. In the current global political environment, a chemical attack would likely be directly or indirectly linked to the atrocities in the theater of the Syrian civil war. Thus, the terrorist use of chemical agents would likely further internationalize this conflict. Security institutions and governments would potentially lose public trust because it would be assumed that they are not able to insulate their societies

effectively from the conflict in Syria and other international threats. In the particular case of phosphine, this connection would be particularly visible because the self-proclaimed Islamic State (ISIS) has been experimenting with the use of phosphine as a chemical weapon in this region (Ackerman and Jacome 2018, p. 29; Binder et al. 2018, p. 28; Quillen 2016, p. 1025; Strack 2017, p. 19).

3. THE STAKEHOLDERS AND THEIR JOINT RESPONSIBILITIES

3.1 The Stakeholders

A variety of stakeholder groups are (or ought to be) involved in the fight against the illegal use of phosphine. However, this chapter will focus on the cooperative actions of three of those groups in particular: security institutions, manufacturers, and vendors of phosphine-based products.

To limit the scope of this analysis, the stakeholders and current counter-measures will be described by using the German counter-terrorism architecture as an example. Other countries, such as the Netherlands or the United Kingdom, show multiple similarities to the German model in their respective security architectures (Van der Veer et al. 2019).

The institutional counter-terrorism architecture in Germany includes a large variety of actors.[3] To improve the communication and cooperation between these actors, a Joint Counter-Terrorism Center (*Gemeinsames Terrorismusabwehrzentrum* [GTAZ]) was established in 2004 that provides a mechanism for the cooperation of a variety of German security agencies concerned with Islamist terrorism (Bundesamt fuer Verfassungsschutz 2017). The GTAZ is not an autonomous institution. It instead acts as a platform to facilitate direct communication between a range of stakeholders in the German counter-terrorism apparatus, such as the police forces, intelligence services, and a variety of ministries.

In addition to the security institutions, the focus of this chapter will be two other groups of relevant stakeholders: the manufacturers of phosphine products and the vendors who sell these products to professional and private customers. These two groups of stakeholders are diverse businesses that interact with each other on local, national, and international levels.

3.2 The Joint Moral Responsibilities of the Stakeholders

All three of the above-described groups have to cooperate to prevent the misuse of phosphine-based products. This cooperation can be portrayed as a multi-layered joint action (see Chapter 3) performed by members of groups

at two levels. The members of each given group perform a joint action with other members of their group (for example, members of the group who sell phosphine act jointly to try to ensure that phosphine is not sold to terrorists, and members of the relevant security agency act jointly to try to ensure that terrorists seeking to acquire phosphine are arrested), giving rise to multiple joint actions. However, this set of joint actions itself constitutes, in turn, a joint action (for example, the joint action performed by members of the security agencies jointly with the manufacturers and the sellers of phosphine, namely, the joint action of preventing phosphine being used in terrorist attacks; Miller 1995, 2001, 2006). Hence, all stakeholders share a joint moral responsibility concerning the prevention of terrorist attacks using phosphine. This joint responsibility manifests itself as a responsibility that is shared between the agents involved in the joint action in question: every single agent is individually responsible for their own contributory, individual action to realize the collective end (that is, the counter-terrorism objective). In virtue of being committed to this collective end, every actor performs their individual action in the belief that the other agents will do the same. In consequence, every agent involved in the joint action is jointly responsible for realizing the collective end – in addition to being individually responsible for their contributory action (Miller 2006).

To be able to perform their individual actions and, thereby, fulfill their respective obligations in accordance with the collective end in question, these individual actions have to be coordinated. This coordination effort becomes even more important if the joint action is a multi-layered one: in the example of the prevention measures against the illegal use of phosphine, every individual in each of the three stakeholder groups performs individual actions as part of a joint action of each group – and that joint action, in return, is part of a higher-order joint action of all three groups together. One option to analyze and coordinate these multi-layered joint actions and responsibilities is the concept of the web of prevention. This concept describes the cooperation of several different (interdependent) groups of agents as well as the interplay of several complementary measures that, together, form a web that helps to prevent the proliferation of certain information and/or materials from different angles and with a certain degree of redundancy (Kuhlau et al. 2012, p. 120; Miller 2018, Ch. 3).

Yet, it still has to be specified what the moral responsibilities of each group in this web are. In the case of the security institutions, this is a rather simple task since the members of these institutions share a moral as well as institutional responsibility to combat terrorist threats against the German Federal Republic. But what of the manufacturers and vendors of phosphine-based products? In the case of these stakeholder groups, the so-called No Means to Harm (NMH) principle can be of help (Miller 2018, Ch. 3). This principle is

used in the academic dual-use debate and states that one ought not to provide others with the means to intentionally inflict large-scale harm on persons or, as I argue, on society. In my interpretation of this principle, psychological or political harm qualifies as large-scale harm to society, at least in some cases.

Equipped with the NMH principle, I argue that manufacturers of phosphine products engage in the production of dual-use substances in the following sense: while a strict prohibition of production of these substances would inflict large-scale harm on the German economy and pest-control efforts, the production of these substances without any considerations of possible malicious use is ethically unsustainable.[4] Hence, manufacturers who are not aware of possible illegal uses of their products ignore their individual responsibility within the web of prevention. Specific threat assessments of their products in cooperation with security institutions are part of their responsibility. Furthermore, manufacturers have a moral responsibility to support efforts to change the composition and design of their products to diminish the dual-use nature of these products.

In addition to the manufacturers, the vendors of phosphine-based products also possess a moral responsibility in this web of prevention based on the NMH principle. However, in contrast to the manufacturers, it is not part of the vendor's responsibility to assess the threats posed by the illegal use of the products that they sell. Rather, vendors only have to be aware that their products might be used in terrorist attacks. However, once the vendors are aware of this security-related relevance of their products, they have to conduct measures to prevent the purchase of said products by malicious agents. The reporting of suspicious purchases is a reasonable way for the vendors to fulfill their responsibilities according to the NMH principle, as described below.

With this web of multi-layered, individual, and joint actions and responsibilities, the common end of preventing terrorists from acquiring phosphine can be realized in an efficacious and ethically sustainable fashion. However, as will be shown in the next section of this chapter, the current cooperative measures to achieve this objective confront multiple problems. I argue that all of these problems stem from the ignorance of the involved stakeholder groups to recognize their respective responsibilities in the web of prevention.

4. CURRENT MEASURES TO COMBAT THE TERRORIST USE OF PHOSPHINE

Due to its toxicity, phosphine is regulated on the German market by different laws. For example, the "Chemikalien-Verbotsverordnung" (*ChemVerbotsV*) and the "Gefahrstoffverordnung" (*GefStoffV*)[5] comprise regulations concerning the purchase of these products by private consumers. The laws allow the purchase of phosphine-producing substances (for example, calcium phos-

phide) only with a permit that requires a government-licensed training in the handling of toxic gases for pest control ("Begasungsschein").

However, there is a worrying loophole in legislation that might be used by malicious agents. Hence, further action is required of both companies and government agencies to fulfill their respective moral obligations. This loophole is an exemption in the *ChemVerbotsV* and *GefStoffV* that states that small amounts of phosphine-producing products are allowed to be sold to private customers without any license. Non-professional customers are allowed to purchase small packages of calcium phosphide tablets without the "Begasungsschein" if the vendor does not suspect any illegal use of the substance.[6] However, the vendor is required to inform the customer about the dangers and possible health hazards connected to the product.

Furthermore, and most important for this chapter, the vendor is required by the *ChemVerbotsV* to document every purchase of phosphine-producing products. Specifically, the identity and address of the customer, as well as the exact amount of purchased products and the intended use of the product by the customer, have to be documented together with the date of the purchase. The records of the purchase have to be archived for at least five years by the vendor. In the German industry and local administration, this documentation is commonly referred to as *Giftbuch* (book of toxins).

With regard to this measure, it has to be noted that government agencies shift the entire responsibility to combat the illegal use of the product to the companies that sell said products. First of all, the documentation of every purchase has to be addressed. The relevant legislation (*ChemVerbotsV* and *GefStoffV*) leaves it open to the vendors in which form they would like to document the purchases of substances like calcium phosphide. This documentation serves two goals: firstly, all purchases of calcium phosphide by customers without a license are documented with the full name and details of the customer so that security agencies can, if necessary, review this documentation and identify suspects or suspicious purchases. Secondly, the documentation of these purchases allows the vendors to limit the number of purchased products to what is called "common amounts for non-professional users" ("haushaltsübliche Mengen").[7]

However, these current measures have at least two problems: current legislation states that non-professional users are allowed to purchase products that produce not more than 15 grams of phosphide per package (see *GefStoffV*, annex 1, 4.2, (2), 2). The purchase of these packages can only be allowed for occasional household use. However, it is not specified what "occasional" means in this regard. Hence, the legislation leaves it to the vendors to decide how the definition of "occasional use" is to be understood in terms of specific amounts of the product. The vendor has to decide in each case if the purchase

of, for example, five packages of calcium phosphide is still a legitimate amount for "occasional use" or already a suspicious purchase that should be reported.

Secondly, it might be impossible to prevent the purchase of excessive amounts of calcium phosphide for vendors in Germany if the customer in question buys only a few packages per store, but in multiple different stores. As confirmed by a representative of a relevant German company, the *Giftbuch* consists of a *physical* notebook in each store that sells calcium phosphide (see above). Therefore, the vendor is not able to detect a purchase of suspicious amounts of this product if this purchase is spread across several stores. It is possible for potential terrorists to purchase large amounts of this product without even disguising their identity.

Lastly, the relevant stakeholders have to be aware of the threat that phosphine poses with regard to terrorism. Since phosphine has not yet been used by terrorists for attacks against Western democracies and is not considered a chemical warfare agent, most of the relevant stakeholders in the web of prevention do not identify this substance as relevant with regard to counter-terrorism.[8]

5. THE WEB OF PREVENTION

The previous section pointed out that the current measures to combat the terrorist use of phosphine offer possibilities for improvement. Yet, these improvements require a certain level of structured cooperation among all stakeholders within the web of prevention. Hence, the foundation of a more efficacious and ethically sustainable collaboration in the web of prevention is the formation of a center in which all relevant stakeholders can meet and communicate directly with each other. An example of how such a center could be organized can be found in Germany with the GTAZ. Other countries, such as the US, have similar approaches, like the National Counterterrorism Center or the Fusion Centers of the Department of Homeland Security (Van der Veer et al. 2019). However, most of these cooperative centers only include one group of stakeholders necessary to form an efficacious web of prevention. While different governmental institutions can communicate and cooperate closely in these centers, relevant businesses are excluded from participating. Yet, as shown in this chapter, only the close cooperation of all mentioned groups of stakeholders contributes to an efficacious and ethically sustainable web of prevention. Hence, a joint center that includes representatives of all these stakeholder groups is needed to prevent the illegal use of phosphine.

5.1 Defining Dangerous Substances

The first point that ought to be discussed by the members of such a joint center is awareness. All groups of stakeholders in the web of prevention are responsible for identifying substances of concern with regard to terrorism or have to be, at least, aware of the relevance of these substances to counter-terrorism efforts. The identification of the danger posed by certain substances shall be the responsibility of the security agencies as well as those companies that manufacture the substances in question. While security agencies possess in-depth knowledge about current trends in terrorism and weapon choices of terrorist groups, manufacturers are naturally aware of the physical properties, the health effects, and the ease of use of their products. Only when both the threat awareness of the security agencies and the technical knowledge of the companies in question are shared is a realistic and efficacious threat assessment for toxic substances possible.

To assess the threat posed by a certain toxic substance, structured and direct communication between security agencies and the manufacturers of phosphine is crucial. Here, the joint center can help to provide a platform where these two groups of stakeholders can meet and share their respective knowledge.

However, as briefly discussed, phosphine does not appear to be a priority of most national security agencies. Moreover, companies that produce products with this substance lack awareness of the possible misuse of these substances. Arguably, one reason for this lack of awareness in security agencies is the fact that phosphine is not expected to cause large amounts of physical damage. However, as previously noted, the impact produced by an attack with these substances is more complex than only the kinetic or health effects. Hence, to assess the complexity of the threat posed by phosphine, the responsible groups of stakeholders should consider using a matrix of threat analysis that also takes psychological and political damage into account.

While security institutions and manufacturers are responsible for jointly assessing the dangers of phosphine, those businesses that sell phosphine products cannot be expected to assess the possible threats posed by this substance. However, as already mentioned, those companies need to be informed about the dangers of the substances that they sell in order to be able to inform security agencies about suspicious purchases. Here, it is the joint responsibility of the security agencies and the manufacturers of these products to share their threat assessment with the vendors that sell them. However, as seen above, German vendors that sell phosphine-producing products are entirely unaware of any security-related issues with the items. To improve awareness in this regard, the joint center could be of help again. The center would offer a simple yet efficacious platform to share the threat assessment with representatives of the vendors who sell products that contain phosphine. As part of the joint

center, the vendors can gain valuable insights into the security-related issues of their products by means of direct communication with the authors of the threat assessment. That would enable this group of stakeholders to be aware of the risks posed by the products they sell and, thereby, to fulfill their moral responsibilities according to the NMH principle.

5.2 What Purchase Is Suspicious?

Once the relevant stakeholder groups in the joint center have identified substances that are considered vulnerable with regard to misuse by terrorists, the participants in the center should discuss ways to deny terrorists access to these substances. As already discussed in the previous section, the stakeholder group of vendors who sell these substances possess a crucial set of responsibilities here. Specifically, the vendors are morally responsible for reporting suspicious purchases of these vulnerable products to security agencies in order to avoid providing others with the means to harm society.

However, to successfully fulfill this moral obligation, the vendors need the other groups of stakeholders as partners in determining what kind of purchases they ought to report to security institutions. Manufacturers and security agencies have to cooperate in defining what one ought to count as a suspicious purchase with regard to phosphine. Since both the manufacturers and the relevant employees of the security institutions share an in-depth knowledge concerning the possible misuses of these substances, they are both capable and responsible for determining what kind of purchases of this substance might be linked to terrorist endeavors. However, this determination has to be shared with the vendors to enable them to fulfill their responsibility and, thereby, to be a functional part of the web of prevention. The joint center, which functions as an organizational application of this web of prevention, offers all three stakeholder groups the forum to define what ought to count as a suspicious purchase for each relevant substance.

Certain toxic substances, including phosphine, are only impactful terrorist weapons if deployed in large amounts. Hence, a purchase might be considered suspicious when judged by the amount of phosphine-producing products that were purchased. In Germany, the definition of what ought to be considered a suspicious (or dangerous) amount of a certain substance is often already provided through legislation.[9] Yet, as seen above, the current system to enforce this legal restriction is not working properly. Hence, the stakeholders at the joint center have to cooperate to find proper detection mechanisms for purchases of suspicious and illegal amounts of dangerous goods.

5.3 How Can We Optimize Detection Mechanisms?

To be able to provide German security institutions with information about the purchases of certain dangerous goods, the relevant vendors are legally obligated to document these purchases in the *Giftbuch*. Yet as already discussed in some detail, this *Giftbuch* is a physical notebook that can be reviewed by local law enforcement or other security institutions if requested. This system of documentation is not sufficient to fulfill the vendor's moral responsibility to actively report suspicious purchases in accordance with the NMH principle. Furthermore, it does not prevent the purchase of an illegal amount of calcium phosphide.

To enable the vendors to live up to their moral responsibilities in the web of prevention and to enforce the existing legislation concerning calcium phosphide purchases, security institutions have to work closely with vendors in the joint center. One possible solution to the current, inadequate measures would be the centralization of the *Giftbuch* in the form of a digital database. By using a cloud-based, digital documentation system, every relevant hardware store employee can check all purchases of calcium phosphide and other dangerous goods that a certain customer made in all connected stores. Equipped with this centralized documentation system, the vendors can easily deny customers excessive amounts of dangerous goods or, if necessary, directly report the customer to the German authorities. German legislators seem sympathetic toward this approach since it is explicitly mentioned in the respective legislation (*ChemVerbotsV*) that the *Giftbuch* can also be present in digital form.[10]

Yet, it would not solve the issues with the current situation in Germany if every relevant vendor would create their own database. To fulfill their moral responsibilities, all relevant vendors have to agree upon an industry-wide documentation system that includes clear rules of access and use by all companies. For example, it might be important to establish rules that prohibit the use of the database for business intelligence-related activities by any party involved. Furthermore, it is crucial that the database fulfills all relevant privacy and data-protection standards of German and European authorities.[11] Lastly, it has to be debated whether the servers for the database should be in possession of and maintained jointly by the relevant vendors, or whether they should be owned by a government institution. The joint center offers the necessary forum for the vendors to discuss these specific issues with each other and with the relevant security institutions.

5.4 Innovating for Security

To conclude this chapter, one additional measure of the stakeholder groups in the web of prevention shall be discussed: the design of technologies for the

societal value of security. By connecting vendors with security institutions and manufacturers, the joint center offers the opportunity to actively discuss the societal value of security while designing new products that might have relevance in this regard. The academic concept of design for values discusses these value considerations in designing new technologies.[12]

Yet, to actively design certain products for the value of security, manufacturers have to, first of all, identify relevant types of products in which designing for security might be fruitful or necessary. Here, again, the joint center can be of help as a forum. Following the NMH principle, it is the joint responsibility of security institutions and manufacturers to discuss the security-related relevance of new and existing products in the joint center. Specifically, the relevant stakeholders ought to identify cooperatively, which domains of products (for example, pesticides) are relevant to counter-terrorism efforts. Subsequently, all members of the joint center ought to discuss the value of security in combination with other societal values that might be (negatively) affected by designing certain products for security.

However, security agencies and manufacturers ought not to be expected to discuss these complex ethical issues all by themselves. The joint center ought to be designed and organized in a value-sensitive manner in order to facilitate the complex debates concerning security and other societal values (Miller 2015). Next to security institutions, manufacturers, and vendors, representatives of citizens and researchers in the fields of applied ethics and social sciences ought to at least be part of the center in advising capacities. As seen above, societal values, such as privacy, autonomy, safety, and security, play pivotal roles in the debates of the groups in the joint center. Hence, competence in applied ethics is needed to steer and moderate these debates. Equipped with this expertise, the stakeholder groups in the joint center can subsequently decide in which way an existing technology or substance ought to be changed or even replaced by a novel innovation to accommodate the value of security. An interesting example of how such a process might look in practice is the case of ammonium nitrate (AN).[13]

Simultaneously with restricting the access to AN fertilizers by means of regulations, European legislators ought to undertake continuous efforts, in cooperation with manufacturers, to change the composition of the substance to make it unattractive for terrorists. For example, the European Union directive 80/876 EEC from 1980 determined that the oil retention of AN prills should not exceed 4 percent. Furthermore, it prescribed that the maximum amount of combustible material in AN fertilizers should not exceed 0.2 percent. Here, it is clearly evident that legislators and manufacturers embedded the values of safety and security into the process of manufacturing AN fertilizers.

Similar efforts to further develop pesticides, such as phosphine, could help to combat the terrorist use of these substances. Designing and redesigning

these products for the value of security ought to be discussed in the joint center and might form an important knot in the web of prevention.[14]

NOTES

1. For example, the Department of Homeland Security considers phosphine a toxic industrial chemical with only moderate risk to be misused as a weapon by terrorist groups or lone operators (TRADOC 2007, Table II-1).
2. The LD_{50} value refers to the lethal dose of a substance and describes how many µg (or mg) per kg body weight of the substance is necessary to kill 50 percent of the exposed population.
3. A more detailed version of this overview was published together with Dr. Paul Burke on https://www.counterterrorismethics.com.
4. For a detailed discussion of dual-use issues in the chemical industry, see Miller and Feltes (2018).
5. The full title of the ChemVerbotsV is: *Verordnung über Verbote und Beschränkungen des Inverkehrbringens und über die Abgabe bestimmter Stoffe, Gemische und Erzeugnisse nach dem Chemikaliengesetz.* The online version is available at: https://www.gesetze-im-internet.de/chemverbotsv_2017/BJNR009410017.html. The full title of the GefStoffV is: *Verordnung zum Schutz vor Gefahrstoffen.* The online version is available at: https://www.gesetze-im-internet.de/gefstoffv_2010/index.html.
6. In the German original, "…wenn, …keine Anhaltspunkte für eine unerlaubte Verwendung oder Weiterveräußerung vorliegen," *ChemVerbotsV*, art. 8, 3, (1). The *GefStoffV* regulates the sale of small amounts of these products in annex 1, 4.2, (2), 2.
7. This is based on an unstructured interview between the author and a representative of one of the most relevant vendors of calcium phosphide in Germany. The interviewee preferred to remain anonymous. The *GefStoffV* refers to this restriction in annex 1, 4.2, (2), 2.
8. This is based on an unstructured interview with a representative of a relevant manufacturer in Germany. The interviewee preferred to remain anonymous.
9. For example, in Germany, customers without a proper license are only allowed to purchase packages of products that contain no more than 15 grams of phosphine (see *GefStoffV*, annex 1, 4.2, (2), 2).
10. *ChemVerbotsV*, art. 9.
11. Note that matters of privacy and data protection are crucial parts in an efficacious and ethically sustainable counterterrorism strategy. Adam Henschke dedicated a chapter to this matter in this volume (Chapter 12).
12. Please note that the research on design for values and value-sensitive design is too extensive to be summarized in this section. Rather, this chapter deploys an applied, general notion of design for security to add to the possible measures that ought to be discussed within the web of prevention (Grunwald 2015; Van den Hoven et al. 2015).
13. AN has been used as the main charge in improvised explosive devices by terrorists. One prominent example is the Oslo bombing in 2011 (Appleton 2014).
14. The analysis in this chapter is derived from Feltes's Ph.D. thesis, submitted to Delft University of Technology and titled *CBRN Threats, Counter-Terrorism and Collective Responsibility.*

REFERENCES

Ackerman, G., and M. Jacome (2018), 'WMD Terrorism', *PRISM*, 7(3), 22–37.
Appleton, C. (2014), 'Lone Wolf Terrorism in Norway', *The International Journal of Human Rights*, 18(2), 127–42.
Binder, M.K., J.M. Quigley, and H.F. Tinsley (2018), 'Islamic State Chemical Weapons: A Case Contained by its Context?', *CTC Sentinel*, 11(3), 27–31.
Bogle, R.G., P. Theron, P. Brooks, P.I. Dargan, and J. Redhead (2006), 'Aluminium Phosphide Poisoning', *Emergency Medicine Journal*, 23(e03), https://doi.org/10.1136/emj.2004.015941.
Bundesamt fuer Verfassungsschutz (2017), *Gemeinsames Terrorismusabwehrzentrum (GTAZ)*, accessed 10 April 2017, at https://www.verfassungsschutz.de/de/arbeitsfelder/af-islamismus-und-islamistischer-terrorismus/gemeinsames-terrorismusabwehrzentrum-gtaz.
Danzig, R., M. Sageman, T. Leighton, L. Hough, H. Yuki, R. Kotani, and Z.M. Hosford (2011), *Aum Shinrikyo. Insights into How Terrorists Develop Biological and Chemical Weapons*, accessed 12 June 2020, at https://www.cnas.org/publications/reports/aum-shinrikyo-insights-into-how-terrorists-develop-biological-and-chemical-weapons.
Grunwald, A. (2015), 'Technology Assessment and Design for Values', in Jeroen van den Hoven, Pieter E. Vermaas, and Ibo van de Poel (eds), *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains*, Dordrecht: Springer, pp. 67–86.
Gurjar, M., A.K. Baronia, A. Azim, and K. Sharma (2011), 'Managing Aluminum Phosphide Poisonings', *Journal of Emergencies, Trauma and Shock*, 4(3), 378.
Gurusinghe, P. (2014), 'Fumigants: Phosphine and Phosphine-Generating Compounds Risk Characterization Document: Environmental Fate', accessed 12 June 2020, at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.441.7484&rep=rep1&type=pdf.
Kuhlau, F., K. Evers, S. Eriksson, and A.T. Höglund (2012), 'Ethical Competence in Dual Use Life Science Research', *Applied Biosafety*, 17(3), 120–27.
Miller, S. (1995), 'Intentions, Ends, and Joint Action', *Philosophical Papers*, 24(1), 51–66.
Miller, S. (2001), *Social Action: A Teleological Account*, Cambridge: Cambridge University Press.
Miller, S. (2006), 'Collective Moral Responsibility: An Individualist Account', *Midwest Studies in Philosophy*, 30(1), 176–93.
Miller, S. (2015), 'Design for Values in Institutions', in Jeroen van den Hoven, Pieter E. Vermaas, and Ibo van de Poel (eds), *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains*, Dordrecht: Springer, pp. 1–11.
Miller, S. (2018), *Dual Use Science and Technology, Ethics and Weapons of Mass Destruction*, Dordrecht: Springer Science & Business Media.
Miller, S., and J. Feltes, (2018), 'Chemical Industry', in Seumas Miller (ed), *Dual Use Science and Technology, Ethics and Weapons of Mass Destruction*, Dordrecht: Springer Science & Business Media, pp. 55–71.
National Center for Biotechnology Information (2019), *Phosphine*, *CID=24404*, accessed 5 April 2021, at https://pubchem.ncbi.nlm.nih.gov/compound/Phosphine.

Pangi, R. (2002), 'Consequence Management in the 1995 Sarin Attacks on the Japanese Subway System', *Studies in Conflict & Terrorism*, 25(6), 421–48.
Parachini, J.V. (2001), 'Comparing Motives and Outcomes of Mass Casualty Terrorism Involving Conventional and Unconventional Weapons', *Studies in Conflict & Terrorism*, 24(5), 389–406.
Quillen, C. (2016, 'The Islamic State's Evolving Chemical Arsenal', *Studies in Conflict & Terrorism*, 39(11), 1019–30, https://doi.org/10.1080/1057610X.2016.1154364.
Smithson, A., and L.-A. Levy (2000), *Ataxia: The Chemical and Biological Terrorism Threat and the US Response* (Report 35), Washington, DC: The Henry L. Stimson Center.
Strack, C. (2017), 'The Evolution of the Islamic State's Chemical Weapons Efforts', *CTC Sentinel*, 10(9), 19–23.
TRADOC (2007), Terrorism and WMD in the Contemporary Operational Environment, *U.S. Army TRADOC G2 Handbook No. 1.04*, Fort Leavenworth, KS: Tradoc Intelligence Support Activity-Threats.
Van den Hoven, J., P. Vermaas, and I. van de Poel (2015), *Handbook of Ethics, Values and Technological Design*, Dordrecht: Springer.
Van der Veer, R., W. Bos, and L. van der Heide (2019), *Fusion Centres in Six European Countries: Emergence, Roles and Challenges*, accessed 12 June 2020, at http://register.consilium.europa.eu/doc/srv?l=EN&f=ST 12168 2005 REV 3.

Index