Towards a Cyber *Jus ad Bellum*: Bridging Legal Gaps within Cyberwar Governance

by

Artur Lukaszczyk

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Philosophy

Waterloo, Ontario, Canada, 2022

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	George R. Lucas, Jr. Professor, Graduate School of Public Policy Naval Postgraduate School
Supervisor	Brian Orend Professor, Department of Philosophy University of Waterloo
Internal Member	Patricia Marino Professor, Department of Philosophy University of Waterloo
Internal Member	Mathieu Doucet Associate Professor, Department of Philosophy University of Waterloo
Internal-External Member	Veronica Kitchen Associate Professor, Department of Political Science University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

While the capabilities of cyberweapons surge forward, our ability to effectively evaluate and govern their deployment has lagged behind. There are presently no internationally binding laws of cyberwar. In their absence, early efforts towards cyberwar governance have revolved around extending existing laws of armed conflict into cyberspace, in hopes of establishing that such laws remained binding within the cyber domain. Although this approach has proven effective at governing cyber operations resulting in physically destructive harms, the decidedly kinetic lens of this approach limits its ability to evaluate the wider spectrum of cyber operations resulting in unfamiliar harms. The goal of this project is to offer a robust evaluative framework which encompasses not only cyber operations resulting in kinetic-analogous consequences, but also disanalogous cyber operations which nonetheless pose a clear and imminent threat to the security of states. I argue that a flexible moral framework built upon the six *jus ad bellum* principles of the Just War Theory tradition offers an avenue by which we may formulate a strong conceptual and ethical foundation for the evaluation of a fuller spectrum of cyber operations, as well as develop norms of best practice for state conduct within cyberspace.

Acknowledgements

I would first like to extend my sincerest gratitude to my supervisor, Dr. Brian Orend, for his continued support throughout my time within the doctoral program. Beginning with a graduate seminar on post-war justice in my first term, to my first focused research area on humanitarian intervention, and finally this dissertation, Dr. Orend's expertise and guidance has proven invaluable throughout my time at the University of Waterloo. From preliminary discussions exploring potential dissertation-worthy topics, to consistently thorough and constructive feedback during the writing process itself, Dr. Orend's unwavering support has made possible the development of a thesis that I am proud of, rather than merely content with. I couldn't have asked for a better supervisor to work with in pursuit of the PhD.

I would also like to express my gratitude to the members of my supervisory committee, Dr. George R. Lucas, Jr., Dr. Veronica Kitchen, Dr. Patricia Marino, and Dr. Mathieu Doucet, for being willing to take the time out of their busy schedules to engage with, and offer guidance for, my project. I am particularly appreciative of Dr. Lucas' willingness to oversee this project as his *Ethics and Cyber Warfare* was the first book I had purchased when exploring potential dissertation topics and drafting a prospective bibliography for what would eventually manifest as this dissertation. As a result, the eventual direction of this project owes much to the early curiosity regarding cyberwarfare that was stoked by Dr. Lucas' work.

I would like to further thank my partner Jessica for her invaluable support throughout my doctoral studies. Words cannot fully express how grateful I am to have had your support throughout the past four years, whether it took the form of fielding my inane questions, reviewing my project ideas, listening to my rants, or simply just dragging me out of the house at the end of a long workday. Thank you for encouraging me throughout the PhD.

I'd further like to express my thanks to Daniel, Oliver, Chris, Liz, Janetta, Piotr, Dean, Michael, and Patrick for helping me keep a reasonable work-life balance in a myriad of ways.

Finally, I would like to thank my family for the time and effort they have put into supporting me throughout this process. I would like to thank my mom, Alina, and my dad, Dariusz, for the sacrifices they have made to provide me the opportunity to undertake this

project. Likewise, I would like to thank my brothers, Adrian and Sebastian, for their support over the past four years, and their efforts to remind me of the world that exists beyond academia.

Table of Contents

List of Figures

Chapter 1

Introduction

Cyberspace is colonising what we used to think of as the real world. I think that our grandchildren will probably regard the distinction we make between what we call the real world and what they think of as simply the world as the quaintest and most incomprehensible thing about us.

William Gibson[1]

1.1 Modern Warfare, Modern Threats

With each passing year, the remaining degrees of separation between cyberspace and "meatspace", the physical world, diminish. Cyberspace has become ubiquitous in contemporary society as the pursuit of progress, efficiency, and convenience has motivated widespread adoptions of network technology. At the personal level, much of our interpersonal socialization now occurs within online contexts, whether through social media platforms or instant messaging services. The banking industry is heavily underpinned by network technologies facilitating international transactions and day-to-day online banking services for clients. Critical government services, such as welfare, have likewise grown more reliant on digital services in the form of online reporting and automated reimbursement. A further emphasis on automation and connectivity has emerged within manufacturing, as manufacturers invest heavily in the development of "smart factories" in the interests of further revolutionizing the industry.

The price of this connectivity is vulnerability. With every new piece of network technology implemented within society, there emerges a new potential target for cyber aggressors. In 2015, attackers using BlackEnergy malware managed to remotely shut down a power grid within Western Ukraine, leaving roughly 230,000 people without power.[2] In 2017, the WannaCry

[1] William Gibson, quoted in Mark Ward, "William Gibson says the future is right here, right now," *BBC News*, October 12, 2010. Accessed June 28, 2022, https://www.bbc.com/news/technology-11502715.
[2] Kevin E. Hemsley and Ronald E. Fisher, "History of Industrial Control System Cyber Incidents," *U.S. Department of Energy*, December 31, 2018, https://doi.org/10.2172/1505628. 16.

ransomware attack significantly impacted the National Health Service within England, infecting the digital systems of numerous hospital trusts, preventing the usage of certain specialized medical equipment, including MRI scanners, and delaying medical treatment for many patients.[3] In 2021, a cyberattack against Colonial Pipeline forced a precautionary shutdown of the pipeline, causing spikes in fuel costs—and very long lines at gas stations—in the southeastern United States.[4] Due to the symbiotic relationship that has developed between meatspace and cyberspace, attacks conducted within cyberspace often spill over into meatspace in some capacity, whether in the form of direct intended consequences and/or as indirect harms caused by disruptions to the cyber systems upon which we have become reliant. Despite originating within cyberspace, cyberattacks have an unignorable effect on the world we inhabit. As a result, it comes as little surprise that cyber warfare has garnered much attention within national defense mandates over the past twenty years.

While our ability to wage war across cyberspace has surged ahead, our ability to evaluate potentially aggressive cyber operations has lagged behind. There are, at present, no internationally binding laws of cyberwar. This absence is made all the more glaring given that our existing laws of armed conflict (LOAC) view war through a decidedly kinetic lens, applying predominantly to those attacks which result in large-scale physical destruction and widespread bodily injury; as such, cyber operations largely fall beyond the purview of the LOAC. International rhetoric regarding the severity of cyber operations proves no more illuminating, as states have asserted that incurred cyber operations *may* be treated as tantamount to armed attacks,[5] while failing to offer any sort of clear threshold beyond which such a designation would be made. As a result, the current paradigm of war within cyberspace is one of uncertainty, within which cyber operators are untethered by conventional governing frameworks of warfare, consistently testing both the defenses and tolerance of their adversaries with progressively more invasive and disruptive operations until an undefined threshold is crossed.

---

[3] S. Ghafur et al., "A Retrospective Impact Analysis of the WannaCry Cyberattack on the NHS," *Npj Digital Medicine* 2, no. 1 (December 2019) https://doi.org/10.1038/s41746-019-0161-6. 1.
[4] Gloria Gonzalez, Ben Lefebvre, and Eric Geller, "Jugular of the U.S. Fuel Pipeline System Shuts Down After Cyberattack," *Politico*, May 8, 2021, accessed May 23, 2021, https://www.politico.com/news/2021/05/08/colonial-pipeline-cyber-attack-485984.
[5] Michael N. Schmitt and Liis Vihul, "The Emergence of International Legal Norms for Cyberconflict," in *Binary Bullets: The Ethics of Cyberwarfare*, eds. Fritz Allhoff, Adam Henschke, & Bradley Jay Strawser (New York, NY: Oxford University Press, 2016), 42.

The goal of this project is to offer a robust evaluative framework for cyberattacks in the form of a cyber *jus ad bellum*. "*Jus ad bellum*" means "the justice of war," and refers to a cluster of concepts and values meant to guide political communities regarding the momentous decision of when to start a war, or take that dramatic first step into large-scale conflict. While formative efforts towards cyberwar governance have sought to extend the LOAC into cyberspace, I argue that these tight-to-the-law evaluative approaches fail to account for the full spectrum of cyber operations. By asserting that *kinetic-equivalent harms* are necessary for invoking the inherent right to self-defense, states targeted by cyber operations remain unable to deploy force in all but the rarest of cases. I argue that the *disanalogous harms* that can be inflicted by cyber measures may pose an equal, if not greater, threat to the functioning of states. Recognizing the limitations of tight-to-the-law evaluative approaches, I argue that a flexible moral framework built upon the *jus ad bellum* principles of the Just War Theory (JWT) tradition offers us a more comprehensive means of evaluating a wider range of cyber operations. This normative approach preserves the right of self-defense even in the event of disanalogous cyberattacks, while nonetheless imposing a set of further constraints to govern permissible and impermissible responses available to a state in the wake of these kinds of cyberattacks.

1.2 The Project Outline

The first substantive chapter, Chapter 2, of this project offers an in-depth review of the contemporary landscape of cyberwarfare. This section begins with a review of the history of cyberwarfare and an exploration of how cyber capabilities have grown to play a significant part in the toolkits of modern militaries and intelligence organizations. This evaluation will include a breakdown of the three primary classes of interstate cyber operations, as well as high profile examples illustrating both the *kinetic* and *non-kinetic* harms these cyberattacks are capable of inflicting. Further attention will be drawn to the notable lack of explicit legislation governing interstate cyber operations and the fraught game-theoretic "equilibrium" we rely on in its absence. I conclude the section by arguing that the stability offered by such norms is insufficient, as poorly defined thresholds of tolerance for cyber operations run the risk of escalations in hostilities stemming from gross discrepancies in how such operations are perceived by states. In

order to dispel these concerns, we require a comprehensive evaluative framework capable of governing uses of force conducted within the grey area of cyberspace.

The third chapter examines a high-profile effort at crafting a cyber-governing evaluative framework, in the form of the *Tallinn Manual*. Drafted by a team of prominent legal experts, the *Tallinn Manual* attempts to extend presently binding international law and legal precedents into the cyber domain as a series of black-letter rules for cyber conduct. This section will examine the rules pertaining to the *use of force* and *armed attack* distinctions within cyberspace, a distinction integral to LOAC evaluations of when multilateral or unilateral forceful responses are permissible. I argue that, while the *Tallinn Manual* offers a strong evaluative framework for cyber operations resulting in familiar *kinetic* harms, it remains limited in its ability to further encompass cyber operations resulting in *disanalogous* harms. Insofar as these latter kinds of operations remain prevalent within the current paradigm of cyber hostilities, we require a more flexible evaluative approach that does not rely on the presence of physical harms. I argue that an extension of the six *jus ad bellum* principles of the JWT tradition offers a strong conceptual and ethical foundation for the development of such an evaluative framework.

The fourth chapter begins developing the cyber *jus ad bellum* framework by first broadening the scope of the *just cause* criterion beyond that within the LOAC, in order to encompass cyberattacks employing novel methodologies or resulting in disanalogous harms. To ground this conception of just cause, this section first analyzes Walzer's important work on aggression and the "common life" residing at the heart of the state and from which the state ultimately derives its rights, including the right to self-defense. I assert that, insofar as cyber operations are equally capable of violating this common life, they may readily constitute just cause even in the absence of inflicting kinetic harms. To establish the thresholds at which the just cause criterion may be met by such operations, I first examine Smith's sovereignty account of just cause, prior to further arguing that cyber operations designed to *manipulate* a state's common life may likewise offer a cyber *casus belli*.

The fifth chapter discusses the remaining two anti-consequentialist principles of *jus ad bellum*, that of right intention and proper declaration, to identify the further constraints that they place on potentially forceful responses to cyber aggression. I assert that the less-inherently-harmful nature of cyber operations renders them more likely to satisfy the right intention

principle than their kinetic cousins. Insofar as both kinetic and cyber responses to aggression may prove equally effective, the more discriminate character of the latter represents a firmer commitment to inflicting no more harm than necessary. I further argue that the principle of public declaration by proper authority enjoys even greater importance within the cyber domain due to cyber operations generally proving more covert in origin and execution than conventional attacks. Given that the potential for mistaken attribution is higher in the event of a cyberattack, the need for public declaration becomes more pressing to prevent potential retaliatory measures against otherwise uninvolved states.

The sixth chapter then moves to cover the three consequentialist principles of *jus ad bellum*: proportionality, last resort, and probability of success. Within this section, I argue that the principle of proportionality heavily restricts the permissibility of kinetic responses to cyber aggression, due to the inherent controversies which accompany comparisons of disanalogous harms. The demands of the last resort principle further restrict the ability of states to respond with conventional force, as the emergence of potent cyber operations considerably broadens the hierarchy of response strategies available to states, bridging severe economic sanctions and lighter conventional military responses with a range of cyber alternatives of escalating severity. I further argue that the probability of success criterion places yet another constraint on the forceful responses available to states as the novel methodology of cyber operations renders it remarkably difficult to conceive of conventional responses likely to be effective at redressing cyber aggression, while still adhering to the remaining principles of *jus ad bellum*. I assert that these three consequentialist principles, taken together, restrict the ability of states to respond with *kinetic force* to solely the most severe cases of cyber aggression.

The seventh chapter illustrates the applicability of the cyber-specific *jus ad bellum* developed throughout the project. To this end, the chapter is comprised of three detailed case studies considering both historical and hypothetical cyber operations. By evaluating these cases through both the LOAC and the cyber *jus ad bellum* frameworks, I argue that the former approach faces significant pragmatic concerns in motivating responses to disanalogous cyberattacks, unfortunately strengthening the position of the aggressor in each case. In contrast, I show how the flexibility afforded by the *jus ad bellum* offers an avenue for states to defend themselves in the event of these three kinds of cyber operations, while nonetheless continuing to minimize the risk of such a response triggering an escalation of hostilities. This, I argue, grants

us a stronger normative foundation from which we may motivate interstate dialogue and strive to develop consensus regarding norms of best practice for cyberwar governance.

Chapter 2

Mapping the Cyber Frontier

2.1 The (Digital) Assassin's Mace

On August 2, 1990, the Iraqi military invaded the neighbouring state of Kuwait following a period of elevated tensions between the two nations. Wracked with a significant amount of debt in the form of loans taken from Kuwait and other nearby states during their conduct of the Iran-Iraq war, the Iraqi government had accused Kuwait of deliberately crippling the Iraqi economy further.[6] Claiming Kuwait had manipulated global oil prices and illegally siphoned oil away from the Rumaila oilfield near the Kuwaiti border, Iraq sought to negotiate a waiver of repayment of the loans it had taken from Kuwait.[7] Upon being rebuffed by Kuwaiti officials, the Iraqi government mobilized its military with the intent of annexing Kuwait. Despite the economic cost of their prior conflict with Iran, the contemporary Iraqi military remained one of the world's largest standing armies, boasting a formidable array of Soviet arms and armour in numbers far exceeding that of their neighbours. As such, the Iraqi offensive made short work of the Kuwaiti resistance, ultimately annexing the state by the end of August 1990.

The Iraqi occupation would last until the early days of 1991, as United Nations Resolution 678 calling for Iraqi withdrawal by January 15th would go ignored by the Iraqi government.[8] Following Iraq's refusal to withdraw, a US-led coalition would launch Operation Desert Storm with the express intent of expelling Iraqi forces from Kuwait in what would become one of the largest mobilizations of conventional military force since WWII. Despite the strength of the battle-hardened Iraqi army, it would prove to be little match for the technologically advanced and numerous combined arms of the coalition forces. The bulk of Iraq's Soviet-era equipment would be destroyed in weeks of air campaigns preceding a brief ground war lasting no more than a hundred hours prior to the declaration of a ceasefire and

[6] Christopher Greenwood, "New World Order or Old? The Invasion of Kuwait and the Rule of Law," *The Modern Law Review* Vol. 55, Issue 2 (March 1992), accessed July 2, 2021: 154-155, https://doi.org/10.1111/j.1468-2230.1992.tb01870.x.
[7] Greenwood, "New World Order or Old?", 155.
[8] United Nations Security Council, *Resolution 678, Iraq/Kuwait*, (29 November 1990).

subsequent negotiations for peace.[9] The casualty lists for both sides would prove similarly lopsided, with estimated Iraqi military deaths ranging between 8,000 to 50,000, a sharp contrast to the coalition casualties numbering roughly 300.[10] Once the dust settled, Desert Storm would be considered a resounding military success, standing as an impressive showcase of the might of the combined conventional arms at the disposal of the coalition forces.

While Desert Storm may have represented the potency of conventional war measures, it also served inadvertently to shift the paradigm of interstate conflict. Following Iraq's emphatic defeat at the hands of coalition forces, the Chinese government, itself sitting on an arsenal of Cold War-era equipment akin to that employed to limited effect by Iraq, began to reconsider its approach to interstate hostilities. To this end, a pair of Chinese People's Liberation Army (PLA) colonels conducted a detailed examination of the United States' military capabilities and their ability to project force globally, ultimately publishing their findings as *Unrestricted Warfare* in 1999.[11] Within their review, the colonels concluded that the United States military would prove insurmountable within the confines of conventional war. As evidenced by the Gulf War, the technological and numerical advantages held by the US military, coupled with the integration of allied forces, ensured they would enjoy unfettered battlefield supremacy in every traditional domain of warfare. Despite the size of the Chinese PLA, it was reliant on arms and armour that were largely outclassed by their Western counterparts. Rather than suggesting that China focus on making up ground on their rivals in the field of traditional weapons, the colonels would instead recommend that China prioritize the proactive development of military capabilities in nascent domains of conflict such as cyberspace.[12]

The resulting doctrine would take on the name of *shashoujian*, or 'assassin's mace', evoking the image of a lightly armed skirmisher overcoming an insurmountable opponent through guile, rather than sheer might. Despite the moniker, *shashoujian* does not itself denote a specific weapon. Rather, the term encompasses the development of tools and tactics for disadvantaged combatants, enabling them to overcome adversaries benefitting from

---

[9] Richard A. Clarke and Robert Knake, *Cyber War: The Next Threat to National Security and What to Do About It* (New York NY: Ecco, 2012), 36.
[10] Britannica, T. Editors of Encyclopaedia. "Persian Gulf War." *Encyclopedia Britannica*, accessed May 6, 2021, https://www.britannica.com/event/Persian-Gulf-War.
[11] George R. Lucas, *Ethics and Cyber Warfare: The Quest for Responsible Security in the Age of Digital Warfare* (New York, NY: Oxford University Press, 2017), 24.
[12] Lucas, *Ethics and Cyber Warfare*, 24.

technological or material disparities which would render them otherwise unassailable.[13] Within the realm of modern warfare, this embodies the usage of non-conventional tactics designed to level the playing field between conventional military superpowers and nations lacking the might to meet them head-on.[14] In the case of China and the PLA, the result is a multi-vector approach to interstate conflict which goes beyond the traditional domains of warfare over which their Western counterparts have typically enjoyed dominance and challenging them in nascent domains within which they have yet to develop a monopoly. Economic warfare-- ranging from targeted sanctions and tariffs to tactical acquisitions of sought-after natural resources-- began to see employment as an alternative means of projecting global influence. Further efforts such as the Chinese "Belt and Road Initiative", which invests heavily into infrastructure projects in developing nations, allow for the Chinese government to develop leverage which may prove beneficial in future negotiations or even conflicts.[15] Likewise, so-called "lawfare", including the pursuit of membership within international legal bodies with intent to shape preferable international policies, offered another means of advancing foreign policy goals without resorting to conventional force.[16] China's push for membership in the UN Human Rights Council illustrates this approach; in June 2020, the Chinese government proposed a resolution suggesting that the current conception of human rights is too often seen as grounds for interference into the sovereignty of other states, offering an alternative form of human rights governance, a change which would hinder the ability of the UN to criticize China's own human rights violations.[17]

Cyberwarfare presents a pivotal third axis for the doctrine of *shashoujian*. In pursuit of warping an adversary's strength into a weakness, the PLA took special interest in both the US' and its allies' sophisticated military equipment as well as their steadily increasing reliance on cyberspace technology.[18] The cyberwar element of *shashoujian* offers what Clarke and Knake

---

[13] Gregory Kulacki, "An Authoritative Source on China's Military Space Strategy," *Union of Concerned Scientists*, March 2014, accessed June 3, 2020: 7, https://www.ucsusa.org/sites/default/files/2019-10/China-s-Military-Space-Strategy.pdf.
[14] Clarke and Knake, *Cyber War*, 37-38.
[15] Lily Kuo and Niko Kommenda, "What is China's *Belt and Road* Initiative?", *The Guardian*, July 30, 2018, accessed July 7, 2021, https://www.theguardian.com/cities/ng-interactive/2018/jul/30/what-china-belt-road-initiative-silk-road-explainer.
[16] Clarke and Knake, *Cyber War*, 37-38.
[17] Sophie Richardson, "China's Influence on the Global Human Rights System," Human Rights Watch, September 14, 2020, accessed July 10, 2021, https://www.hrw.org/news/2020/09/14/chinas-influence-global-human-rights-system.
[18] Lucas, *Ethics and Cyber Warfare*, 24.

identify as two specific 'paths of improvement' for the PLA. The first of these is a strategy of theft; cyberattacks targeting sensitive information pertaining to an adversary's military technologies enable the discovery of weaponizable exploits for use in future conflicts, as well as making it possible to adapt the same technologies for the modernization of one's own military.[19] This approach offers a bipartite boon for Chinese military strategy, enabling the PLA to rapidly bridge the technological gap between it and Western states, while simultaneously accelerating the development of tactics designed to hinder, or altogether eliminate, their adversaries' technological superiorities. In pursuing this strategy, even US aircraft carrier battle groups, the crown jewel of the US navy and a lynchpin in the US' ability to project military power globally, would be discovered to have technological vulnerabilities: e.g., following the deployment of a pair of US carrier battlegroups to Taiwan in 1996 due to heightened tensions between Taiwan and China, Chinese Air Force officers would publicly reveal that the battlegroups' effectiveness could be significantly reduced by cyberattacks against their information systems.[20] The second strategy is one of avoiding conventional confrontation entirely, instead relying on the crippling potential of cyberattacks directed against nations heavily reliant on network technologies.[21] The ability to remotely target everything from military command-and-control centers to civilian digital infrastructure can serve to reduce not only an adversary's ability to coordinate a response and fight, but also their will to continue either their aggression or their resistance. Accordingly, an offensive cyber strategy offers a potent tool for rendering potential adversaries more amicable to favourable resolutions, without the need to resort to conventional armed conflict. Such would indeed serve, to quote the legendary Sun Tzu, as a novel, clever and formidable "art of war."

In the spirit of *shashoujian*, the Chinese government has developed two cyber-specific branches of the PLA in the form of Units 61398 and 78020.[22] Since their inception, both units have regularly engaged in a myriad of cybercrimes seeking to advance foreign policy in the digital domain. These cyberattacks would include the theft of industrial technologies as well as sabotaging civilian infrastructure through the use of "logic bombs" and other digital weapons.[23] Further attacks would be conducted against various research institutes and universities in hopes

[19] Clarke and Knake, *Cyber War*, 39.
[20] Clarke and Knake, *Cyber War*, 39.
[21] Clarke and Knake, *Cyber War*, 39.
[22] Lucas, *Ethics and Cyber Warfare*, 24.
[23] Lucas, *Ethics and Cyber Warfare*, 24-25.

of gleaning cutting-edge technologies being developed by their researchers. Military technology programs, both government projects and joint efforts between the military and defense contractors, have proven to be targets of particular interest for China-based cyber operatives insofar as attacks against them could serve to simultaneously unearth the capabilities and limitations of sensitive foreign military equipment, while accelerating the procurement efforts of Chinese manufacturers. Accordingly, cyberwar tactics have evolved into a key component of China's foreign policy, allowing them to advance their foreign policy goals and project their influence globally, without having to resort to the employment of conventional force which may draw them into an undesirable engagement with the traditionally dominant military powers.

The integration of cyber elements into foreign policy strategy is by no means unique to China and the PLA. Much like China, Russia does not identify cyberwarfare as a standalone concept. Rather, cyber operations are conceptualized as merely a component of the grander scheme of *informatsionnaya voyna*, or information warfare.[24] Information warfare represents Russia's strategy for controlling the information landscape through the combined deployment of tactics ranging from electronic warfare to psychological operations.[25] At their core, information warfare operations seek to undermine the foreign policy goals of adversaries, as well as their ability to hinder or resist the advancement of one's own goals, without resorting to conventional military means or physical force.[26] This style of warfare represents a key tool for Russian foreign policy, as Russian military strategists have deduced that dominance within the information sphere can be leveraged to undermine the legitimacy of foreign governments and sow discord amongst their constituents, increasing the effectiveness of Russian efforts abroad without the need for protracted conventional conflict.[27] Pivotally, contemporary Russian military doctrine does not conceive of information warfare as merely another wartime tool to be used in conjunction with conventional force; rather, it specifically states that modern conflicts are themselves preceded by "*the prior implementation of measures of information warfare* in order to achieve political objectives without the utilization of military force and, subsequently, in the

[24] Michael Connell and Sarah Vogler, "Russia's Approach to Cyber Warfare," *CNA* (March 2017), accessed June 2, 2021: 3, https://www.cna.org/cna_files/pdf/DOP-2016-U-014231-1Rev.pdf.
[25] Connell and Vogler, "Russia's Approach to Cyber Warfare," 3.
[26] Ofer Fridman, "Information War as the Russian Conceptualisation of Strategic Communications," *The RUSI Journal* Vol. 165. Issue 1. (March 2020), accessed June 5, 2021: 47, https://doi.org/10.1080/03071847.2020.1740494.
[27] Connell & Vogler, "Russia's Approach to Cyber Warfare," 4.

interest of shaping a favourable response from the world community to the utilization of military force".[28] Consequently, Russia's conception of information warfare clearly demarcates it as a measure short-of-war which may justifiably be employed as an alternative means of achieving Russian foreign policy objectives, even in times of peace.

While the introduction of cyber warfare may not have entirely revolutionized the Russian concept of *informatsionnaya voyna*, it has undoubtedly served to expand the information warfare playbook. An increasingly networked world offers unique mediums for waging information warfare. Whereas past efforts at controlling the narrative of conflict were necessarily conducted through low-tech means such as state-sanctioned news networks broadcasting propaganda, the proliferation of communications technology has expanded the reach of information warfare operators. Today, news has become increasingly decentralized as bloggers, online discussion forums, and social media sites skew public perception on a myriad of topics and global events. Consequently, information warfare operators now find themselves capable of manipulating public perception across multiple spheres of influence; sweeping state-sanctioned news reports can paint broad misleading narratives, while tailored appeals to targeted demographics via social media manipulation can shape opinions at a more intimate level. Likewise, further floods of disinformation sowed via bot contributions to discussion boards can serve to muddy discourse and proliferate confusion. Crucially, the covert nature of cyber operations carries with it the invaluable element of plausible deniability; the use of botnets, aliases, and proxies allows information warfare operators to manipulate the information landscape while disguising their efforts to appear as though they have manifested organically, rather than being carefully crafted narratives designed to advance specific foreign policy goals.[29]

The effectiveness of information warfare strategies is perhaps best evidenced by the Russian campaign within Ukraine following the ousting of pro-Russian Ukrainian President Viktor Yanukovych, ultimately culminating with the annexation of the Crimean Peninsula. In the wake of domestic unrest within Ukraine following the 2013 pro-European Euromaidan protests in the capital city of Kyiv, a Russian-linked hacktivist group named CyberBerkut launched a

[28] Connell and Vogler, "Russia's Approach to Cyber Warfare," 3.
[29] Connell and Vogler, "Russia's Approach to Cyber Warfare," 4-5.

series of cyberattacks aimed at destabilizing the new, post-Yanukovych Ukrainian government.[30] The group, having adopted the moniker in reference to the infamous Ukrainian Berkut special police which had worked to suppress the Euromaidan protests, would carry on their attempts to shut down pro-European sentiment amongst Ukrainians, albeit within the digital domain. Their campaign would be comprised of varying tactics, ranging from website defacements designed to paint the new Ukrainian government as fascists, to Distributed Denial of Service (DDoS) attacks aimed at disrupting the interim government's ability to coordinate a response and de-escalate the situation.[31] Further efforts were made to skew public perception regarding Ukraine's subsequent domestic elections for Yanukovich's successor, with CyberBerkut operatives going on to sabotage Ukraine's central election computers and planting malware designed to declare an election victory for the ultranationalist party, further propagating the narrative that the uprisings were a fascist ploy.[32] These cyber tactics were employed to amplify the rift in Ukraine's domestic politics, seeking to create a narrative of domestic support for maintaining strong ties with the Russian Federation, while simultaneously downplaying pro-European sentiment as a kind of inorganic fabrication orchestrated by Western powers.

Capitalizing on the generalized chaos it had caused within Ukraine's domestic politics, Russia's information warfare campaign would also pave the way for the annexation of the Crimean Peninsula into the Russian Federation. As the new Ukrainian government worked to restore order in the post-Yanukovych era, Russia would turn their attentions south. Under the guise of protecting ethnic Russians residing within Crimea from the new "illegitimate" government (roughly 58.5% of Crimean residents had identified as ethnically Russian in a 2001 Ukraine-wide census),[33] Russia would supplement their information warfare efforts kinetically with the deployment of 'little green men' (presumed to be unmarked Russian special forces troops) in late February of 2014. These troops would proceed to seize numerous government and military sites in the region, seriously subverting the Ukrainian government's ability to govern

[30] Benjamin Jensen, Brandon Valeriano, and Ryan Maness, "Fancy Bears and Digital Trolls: Cyber Strategy with a Russian Twist," *Journal of Strategic Studies* Vol. 42. Issue 2. (January 2019), accessed May 29, 2021: 13-14 https://doi.org/10.1080/01402390.2018.1559152.
[31] Jensen, Valeriano, and Maness, "Fancy Bears and Digital Trolls," 13-14.
[32] Jensen, Valeriano, and Maness, "Fancy Bears and Digital Trolls," 14-15.
[33] John Biersack and Shannon O'Lear, "The Geopolitics of Russia's Annexation of Crimea: Narratives, Identity, Silences, and Energy," *Eurasian Geography and Economics* Vol. 55. Issue 3. (December 2014), accessed July 11, 2021: 254, https://doi.org/10.1080/15387216.2014.985241.

within the peninsula. With Crimean irredentist aspirations stoked by concerted disinformation campaigns and the central government's authority subverted by the presence of well-armed troops, the de facto government of Crimea held a controversial referendum in early March. The referendum would offer a choice between Crimea entering the Russian Federation or increased regional autonomy, reducing the central Ukrainian government's power over the region.[34] The referendum would show overwhelming support for the former option with 96.77% of the votes expressing support for reunification with Russia.[35] Despite concerns regarding voting irregularities and the circumstances surrounding the referendum vote, Crimea would be formally annexed into the Russian Federation in mid-March. Serious international sanctions on Russia were, eventually, levelled in response to this action, yet it remains a done deal, on the ground, still today. Amongst other things, it reveals: a) the potent capability of blended cyber-kinetic strategies; and b) how some of the very oldest objectives of war, such as territorial expansion, can be advanced through some of the very newest cyber means.

The burgeoning emphasis on cyber capabilities by their traditional adversaries would not go unnoticed by the United States and other NATO members. Since at least the early 2000s, concerns regarding cyberwarfare have begun to appear more frequently in US Department of Defense (DoD) and Department of Homeland Security (DHS) briefings as foreign cyber strategies began to develop. In response, initial US forays into cyber strategies of their own were predominantly limited to the establishment of deterrent policies designed to dissuade foreign cyberattackers from targeting American interests. This would prove short-lived as US cyber policy would eventually evolve to task US Cyber Command (USCYBERCOM) with the development of offensive cyber capabilities which could, in turn, be used as a deterrent force.[36] This shift from defensive to offensive posturing is motivated by an acknowledgement that cyber defenses are often playing catchup to offensive cyberweapons. Many cyber weapons are designed to take advantage of undetected vulnerabilities, such as so-called "zero-day exploits" which are software vulnerabilities capitalized on by attackers prior to the vulnerability being discovered and patched by the developer.[37] This renders cyber defense an incredibly difficult,

[34] Biersack and O'Lear, "The Geopolitics of Russia's Annexation of Crimea," 251.
[35] Biersack and O'Lear, "The Geopolitics of Russia's Annexation of Crimea," 251.
[36] Alex S. Wilner, "US Cyber Deterrence: Practice Guiding Theory," *Journal of Strategic Studies* Vol. 43. Issue 2. (February 2019), accessed June 5, 2021: 260-261, https://doi.org/10.1080/01402390.2018.1563779.
[37] "What Is a Zero-Day Exploit?", FireEye, accessed July 11, 2021. https://www.fireeye.com/current-threats/what-is-a-zero-day-exploit.html.

time consuming, and cost-intensive task. Recognizing the acute asymmetry in capability existing between offensive and defensive cyber action, US cyber-strategy analysts have cautioned that the US' cyber defense capabilities will likely pale in the face of foreign cyber weapons for at least the near future.[38] Accordingly, the best interim cyber defense needs to be a potent cyber offense.

The 2010 Stuxnet worm represented a key evolutionary step in the development of these cyber offensive capabilities. Believed to be the result of a collaboration between Israel and the United States, Stuxnet was a cyber weapon designed as a means of disrupting Iran's controversial nuclear program through remotely sabotaging a nuclear plant in Natanz.[39] Despite the control systems for the plant being "air-gapped", isolated from potentially unsecured networks and the internet in the interests of security, the Stuxnet worm would eventually infiltrate the systems responsible for controlling the plant's nuclear centrifuges by infecting and travelling across the USB sticks of Iranian engineers.[40] Having made its way into the centrifuge array, Stuxnet would then seize control of the systems, feeding the Iranian operators false data showing that all systems were operating normally, while secretly causing the arrays to behave in an unsafe manner, ultimately leading them to self-destruct.[41] The deployment of Stuxnet managed to set Iran's nuclear program back significantly, without the additional collateral damage that would inevitably ensue in the event of a conventional alternative such as an airstrike. Accordingly, Stuxnet stands as proof that cyberweapons have the potential to achieve effects that were previously the sole domain of conventional attacks.

Evidently, cyberwarfare has grown from relative obscurity to a full-fledged domain of conflict. At present, nearly every developed nation has established cyber-specific military units trained for conflict within the cyber domain. Likewise, cyber operations have evolved from being an afterthought to an integral part of modern military strategy; the focus of defense policies now shifts away from the conventional wars and mobilizations of force that defined the 20th century, and towards low-visibility, high-impact digital operations which can in many cases achieve the same objectives as conventional force can, albeit at a lower financial cost and with a

[38] Wilner, "US Cyber Deterrence: Practice Guiding Theory," 263.
[39] Brian Orend, *War and Political Theory.* (Cambridge, UK; Medford, MA: Polity, 2019), 176.
[40] Andy Greenberg, *Sandworm: A New Era of Cyberwar and the Hunt for the Kremlin's Most Dangerous Hackers*. (New York, NY: Doubleday, 2019), 116.
[41] George R. Lucas Jr., "Emerging Norms for Cyber Warfare," In *Binary Bullets: The Ethics of Cyberwarfare*, eds. Fritz Allhoff, Adam Henschke, & Bradley Jay Strawser (New York, NY: Oxford University Press, 2016), 27.

decreased risk of kinetic retaliation. Given the advantages cyberwarfare offers to nations unable or unwilling to commit to open warfare, it is likely that the early 21st century will mark a sort of digital arms race as each nation seeks to establish themselves as a force within the nascent cyber domain over which no nation has yet staked a claim to supremacy. This is made all the more likely as superpowers which have traditionally enjoyed unfettered military supremacy, such as the United States, now find themselves vulnerable. And insecurity, quite often, can produce not merely re-armament but actual conflict.

2.2 The Anatomy of a Cyberattack

As world powers gear up for conflict within the "fifth domain", so too has the frequency of cyberattacks increased. Presently, thousands of cyberattacks occur daily, targeting everything from personal devices to government and corporate databases. Most of these attacks represent comparatively minute threats, such as phishing attempts targeting civilian email accounts. Nonetheless, more severe cyberattacks posing further reaching and more impactful harms have begun to grow in prevalence over the last decade. Cyberattacks at the end of 2015 were responsible for remotely shutting down a power grid in Western Ukraine in the first known case of a hacker-induced blackout.[42] In 2017, the "WannaCry" cyberattacks originating from within North Korea resulted in major disruptions for the British National Health Service as hospital computers, MRI scanners, and other medical equipment were infected by ransomware, rendering many NHS patients unable to receive care.[43] In 2020, multiple US government agencies including the Department of Homeland Security (DHS) fell victim to the SolarWinds hacks, believed to be a large-scale data theft operation perpetrated by Russian intelligence.[44] These SolarWinds hacks also functioned as a formidable "show of force", so to speak, of Russia's cyber-capability: striking deep into the heart of American governing agencies. On top of the actual attacks, there is the potential for further gains in the form of geo-political intimidation and manipulation which successful cyberstrikes can enable. Even more recently, a ransomware attack

[42] Greenberg, *Sandworm*, 61-62.
[43] Orend, *War and Political Theory*, 173.
[44] David E. Sanger, Nicole Perlroth, and Eric Schmitt, "Scope of Russian Hack Becomes Clear: Multiple U.S. Agencies Were Hit," *New York Times*, December 14, 2020, accessed May 23, 2021. https://www.nytimes.com/2020/12/14/us/politics/russia-hack-nsa-homeland-security-pentagon.html.

on the US-based Colonial Pipeline in 2021 forced the pipeline to shut down after hackers seized swathes of company data, causing a run on gasoline supplies by ordinary Americans, and drawing global attention to the risks posed by cyberattacks to aging critical infrastructure fraught with digital vulnerabilities.[45]

While cyberattacks are understandably varied in their scope and approach, they can be broken into three overarching categories: espionage, disinformation, and sabotage.[46] Cyber espionage involves the usage of computer technologies to access and steal sensitive information.[47] This often takes the form of spear phishing attacks which target specific individuals, often high-ranking members of government or corporations, with disguised emails designed either to trick the target into revealing sensitive information, or to deploy malware tasked with extracting sensitive intelligence directly. Unlike regular phishing campaigns, within which misleading emails are generically phrased and indiscriminately disseminated, spear phishing attacks represent concerted efforts to gain access to specific government organizations or corporations. The 2016 hacks into the Democratic National Committee were partly comprised of spear phishing attacks targeting high ranking members of the DNC with the intent of monitoring their communications in the lead up to the US election; it was ultimately one such spear phishing attack which would prove successful in accessing Democratic Campaign Chairman John Podesta's emails, allowing them to be extricated and subsequently released via WikiLeaks, generating controversy and causing the election campaign to fall into some disarray.[48]

The targets of cyber espionage can vary dramatically, as can the intent behind the attacks. The Chinese doctrine of *shashoujian* has resulted in a prolific campaign of cyber espionage, with hacks attributed to operatives within China having hit numerous high-profile targets throughout the 2000s. Among these targets are global corporations such as Coca-Cola, targeted by cyber operatives in hope of unearthing invaluable trade secrets which can be used to bolster domestic

[45] Gloria Gonzalez, Ben Lefebvre, and Eric Geller, "Jugular of the U.S. Fuel Pipeline System Shuts Down After Cyberattack," *Politico*, May 8, 2021, accessed May 23, 2021, https://www.politico.com/news/2021/05/08/colonial-pipeline-cyber-attack-485984.
[46] Orend, *War and Political Theory*, 173-174.
[47] Orend, *War and Political Theory*, 173-174.
[48] Jensen, Valeriano, and Maness, "Fancy Bears and Digital Trolls," 9.

corporations, further contributing to China's broader campaign of economic growth/warfare.[49] Further attacks have targeted the US military and government-adjacent defense contractors to access sensitive information regarding next generation military technologies to be repurposed for China's own modernization effort, such as their domestic stealth fighter program.[50] Perhaps more insidiously, cyber espionage has also been employed extensively for political purposes. Chinese operators have made use of cyber espionage tactics to collect data regarding political targets of interest abroad; the Chinese GhostNet operation in 2009, which had penetrated India's Foreign Ministry and Ministry of Defense, was designed to collect information regarding the exiled Tibetan community residing within India.[51] As evidenced by the DNC hacks, cyber espionage has likewise offered Russian intelligence services a new means of gathering *kompromat*, or compromising material, pertaining to targets of interest. Finally, cyberespionage has likewise served as an extension of traditional military espionage, as Russian intelligence proved able to uncover the positions of Ukrainian artillery units by infecting unsecured apps used by Ukrainian troops for targeting calculations, bolstering the effectiveness of their conventional war efforts during the ongoing conflict in Ukraine.[52]

This brings us to the second form of cyberattack. Disinformation represents the deliberate manipulation or obfuscation of information via computer technologies, designed to undermine a target's core interests.[53] Cyber disinformation exists within two specific dimensions. The first of these is *strategic disinformation* within which disinformation techniques are used to disorient an adversary's defensive capabilities for the purposes of enabling more conventional kinetic action, be they air strikes or the deployment of special forces. This may be comprised of cyberattacks designed to mask the movements of one's own troops, or to disorient an adversary regarding the status of theirs. This form of disinformation was utilized to great effect by the Israeli Air Force (IAF) as a precursor to a bombing raid on a suspected weapons facility in Diaya-al-Sahir in 2007. In hopes of masking the flight path of IAF bombers, the Israelis managed to remotely disable Syrian anti-air defenses, subsequently feeding Syrian air space observers false data

[49] Center for Strategic & International Studies, "Significant Cyber Incidents," accessed May 15, 2021: 55, https://www.csis.org/programs/strategic-technologies-program/significant-cyber-incidents.
[50] Center for Strategic & International Studies, "Significant Cyber Incidents," 49.
[51] James P. Farwell and Rafal Rohozinski, "Stuxnet and the Future of Cyber War," *Survival*, Vol. 53. Issue 1. (January 2011), accessed May 27, 2021: 26, https://doi.org/10.1080/00396338.2011.555586.
[52] Jensen, Valeriano, & Maness, "Fancy Bears and Digital Trolls," 16.
[53] Orend, *War and Political Theory*, 174.

hiding the IAF incursion into their territory.[54] Ultimately, the raid proved successful, the undetected bombers managing to destroy the target facility without the risk of lost airplanes and casualties. In this case, strategic disinformation was used to undermine a target's interests by hamstringing their ability to defend their territory through computer technologies designed to undermine their defensive systems. Notably, strategic disinformation represents the use of cyberattacks as a tool employed in tandem with conventional warfare tactics as a "force multiplier", rendering the latter more effective.[55]

On the other hand, *narrative disinformation* is less concerned with duping military early-warning radars and air defenses. Rather, narrative disinformation aims to advance foreign policy goals through the careful control, distortion, and manipulation of information abroad. This approach may manifest in various forms, including the purposeful dissemination of false information to a targeted populace designed to inflame political tensions and paralyze domestic politics. At an extreme, it may constitute concerted propaganda efforts geared towards fanning the flames of irredentist aspirations and triggering a revolution within an adversary's borders. Narrative disinformation may also be employed to shape public perception regarding interstate conflict, either to garner widespread support for one's own conventional military actions or to trigger public outcry against those of an adversary. Whereas strategic disinformation typically targets the state directly, and is more short-term, narrative disinformation is often more sweeping in scope, operating over a longer-term through the mass influencing of public perception to foster favourable sentiments.

This narrative form of disinformation is regularly employed by Russian intelligence in the conduct of information warfare. In the case of Ukraine, the Russian disinformation efforts were launched following concerns that the Ukrainian government would resume their campaign towards gaining both EU and NATO membership for the country, causing the country to drift further from the Russian sphere of influence.[56] Following the exile of Yanukovich to Russia, extensive narrative disinformation attacks were conducted by CyberBerkut, presumed to be linked with another hacking group designated Fancy Bear; the latter Russian-based group is

[54] Lucas, "Emerging Norms for Cyber Warfare," 27.
[55] Orend, *War and Political Theory,* 173.
[56] Greenberg, *Sandworm*, 58.

widely believed to itself be a cyber unit affiliated with Russia's military intelligence, the GRU.[57] Hacktivists launched a variety of cyberattacks ranging from the spread of pro-Russian propaganda to the disruption of the communications of pro-Western revolutionaries, each designed to paint the ongoing situation in Ukraine as a Western-controlled junta seeking legitimacy through a sham election.[58] Fancy Bear later conducted its own narrative disinformation campaign in the lead-up to the 2016 US Presidential Election, acquiring sensitive emails from both the DNC and RNC, releasing the emails of the former to distort electoral discourse in accordance with Russian foreign interests. In a retrospective, the US Intelligence Community concluded that Russian disinformation efforts had been specifically designed to "undermine public faith in the US democratic process, denigrate Secretary [Hillary] Clinton, and harm her electability and potential presidency".[59] In both cases, narrative disinformation attacks were able to severely impact the domestic politics of a targeted foreign state (and a very formidable and well-endowed one, at that) without resorting to openly combative measures. Consequently, disinformation attacks offer cyberpowers a dual-purpose tool geared towards dominance in both conventional and unconventional forms of interstate conflict.

The third category of cyberattacks is that of sabotage. Cyber sabotage entails the usage of computer technology to either impair or destroy systems necessary for a political community's core interests.[60] Presently, almost every facet of society exhibits heavy reliance on computer technologies, with more industries and services shifting towards having some online presence each year. At an individual level, interpersonal communication and access to information is predominantly facilitated by internet access. Financial corporations such as banks and insurance companies now conduct much of their business digitally. Public health services are likewise dependent on interconnected databases for patient records and the transfer of sensitive medical data, while many diagnostic machines are similarly connected to wireless hospital networks. Programmable logic controllers (PLCs) underpin much of the efficiency of industries such as oil and manufacturing. The effectiveness of modern militaries is likewise built upon the strength of their informatics systems, as targeting and information control systems grow ever more complex. Even core municipal services such as water filtration and power grids are often controlled via

[57] Greenberg, *Sandworm*, 58.
[58] Greenberg, *Sandworm*, 57.
[59] Jensen, Valeriano, and Maness, "Fancy Bears and Digital Trolls," 222.
[60] Orend, *War and Political Theory*, 174.

computer technologies with limited human oversight. And the COVID-19 pandemic, of course, has only widened and deepened such technological trends.

While the benefits of integrating computer technologies are readily apparent, this shift also renders these systems vulnerable to acts of cyber sabotage. The widespread cyberattacks on Ukraine showed that digital infrastructure can be shut down by belligerents. The Stuxnet worm further evidenced how cyberattacks may easily transcend the digital-physical divide and result in real-world harms. These past high-profile attacks have led to vociferous doomsayers pointing to a potential future of cyber-sabotage regularly resulting in high-profile disasters, such as refinery fires brought about by remotely manipulated PLCs tampering with the flow of oil.[61] While cataclysmic cyberattacks of this degree seem to pose a more remote threat, there is nonetheless concern regarding the effects of smaller scale physical harms brought about by cyberattacks against unguarded infrastructure. One such attack occurred against a water treatment facility for the city of Oldsmar, Florida in 2021. Having gained access to the control interface of the facility after a plant worker accessed a malware-ridden site, hackers were able to remotely change the amount of lye being introduced into the water system from 100 parts per million to 11,100, rendering it potentially fatal to ingest.[62] Fortunately, the change was noticed by supervising workers at the plant and reversed before any real damage could occur. Nonetheless, these sorts of low-profile cyberattacks present a very real risk for more rudimentary infrastructure, particularly given that less emphasis has hitherto been placed on "ruggedizing" these technologies against digital threats when compared to higher-profile potential targets such as national power grids and government databases.

Notably, acts of cyber sabotage need not necessarily cause physical damages, nor do they need to result in permanent harms. In many cases, cyber sabotage manifests as widescale temporary disruptions to a state's critical infrastructure. Botnet-driven offensives, such as DDoS attacks, are capable of flooding webpages with enough traffic to knock them offline, hindering the provision of important services such as online banking. This sort of attack hit Estonia in 2007 following a controversial decision by the Estonian government to remove a Soviet-era military

[61] Steven P. Lee, "The Ethics of Cyberattack," In *The Ethics of Information Warfare*, eds. Luciano Floridi & Mariarosaria Taddeo (Heidelberg: Springer International Publishing, 2014), 108.
[62] Dan Goodin, "Florida Water Plant Compromise Came Hours After Worker Visited Malicious Site," *Ars Technica*, May 18, 2021, accessed May 23, 2021, https://arstechnica.com/gadgets/2021/05/florida-water-plant-compromise-came-hours-after-worker-visited-malicious-site/.

statue from the center of the capital city Tallinn. Following severe backlash from both the Russian government and ethnic Russians within Estonia, the country was struck with blanket DDoS attacks which shut down everything from banks and hospitals, and even niche domestic community forums.[63] Despite Estonian accusations, the Russian government subsequently denied any involvement in the attacks, ultimately claiming that "[Russia] can't be blamed if individual Patriots take matters into their own hands".[64] Another act of widespread cyber sabotage would come a decade later in the form of NotPetya, a worm disguised as ransomware which ravaged Ukraine in 2017. NotPetya served to sabotage nearly every aspect of Ukrainian infrastructure, shutting down everything from the Ministry of Health to the post office, prompting Ukraine's minister of infrastructure to declare that "[t]he government was dead".[65] Neither attack resulted in the victimized nation suffering physical or irreversible harms. Nonetheless, in both cases, cyber sabotage proved more than capable of significantly disrupting a targeted state's core interests.

While cyberattacks may be divided into these three categories, it regularly proves to be the case that examples of cyberattacks fail to fall neatly into one of the three. Many times, a cyberattack manifests as a hybrid of two or three categories, with elements of one type of attack proving conducive to the facilitation of another. For example, espionage attacks may readily evolve into disinformation operations as the stolen information is weaponized and subsequently used to manipulate the information landscape. The Russian disinformation campaign around the 2016 US Presidential Election was preceded by comprehensive espionage efforts seeking to gain access into the playbooks of both the DNC and its Republican counterpart, later selectively releasing sensitive information lifted from the DNC to help shape the electoral narrative. Similarly, disinformation attacks often work in tandem with cyber sabotage as the latter may be employed to hinder an adversary's ability to respond to disinformation, amplifying the impact of the former. These kinds of attacks naturally prove to be a potent combination when paired with conventional military tactics. This type of hybrid warfare emerged within Russia's 2008 invasion of the neighbouring country of Georgia, during which Russian cyber operatives launched sabotage attacks against Georgian political offices and local media, while Russian-affiliated

[63] Greenberg, *Sandworm*, 96.
[64] Lucas, "Emerging Norms for Cyber Warfare," 26.
[65] Greenberg, *Sandworm*, 197.

media pushed falsified reports of Georgian troops massacring civilians in South Ossetia and Abkhazia.[66] Not only did these cyberattacks lay the groundwork for Russia's declaration of its invasion of Georgia under the pretense of protecting South Ossetia and Abkhazia, but the acts of cyber sabotage would continue throughout the hostilities with attacks against power grids causing generalized chaos while persistent attacks against Georgia's government infrastructure ensured that any response mustered by Georgian forces would be fragmented at best. It was a potent example of contemporary hybrid, or "two-track", cyber-physical warfare.[67]

Evidently, the cyber domain is one of growing importance in the field of interstate conflict. Nearly every facet of modern society is now underpinned by digital components, ensuring that cyber operatives are spoiled for choice when it comes to target selection. The shift away from conventional military force has likewise proven a boon for smaller nations unable to survive a protracted conventional conflict due to the sheer scale of conventional military strength that nations such as the United States and Russia can bring to bear. For these smaller global actors, cyberwar offers a low-cost, high impact means of disrupting the machinations of larger states, given that cyber defenses are often much harder to formulate than offensive measures. Acknowledging this asymmetry, larger nations have likewise pursued the development of cutting-edge offensive measures designed to serve as a deterrent while establishing themselves as leading global cyberpowers. Accordingly, cyberspace has become an experimental domain within which each nation procures and employs their own cyber weapons as a supplement to their foreign policy- and national defense strategies.

2.3 Shots in the Dark: When is Cyberwar "War"?

In a 2013 article discussing the changing character of warfare, General Valery Gerasimov, Chief of the General Staff of the Russian Armed Forces, noted that the 21st century has been marked by a "tendency toward blurring the lines between the states of war and peace. Wars are no longer declared and, having begun, proceed according to an unfamiliar template".[68] Perhaps nowhere do Gerasimov's observations hold more weight than within the cyber domain.

[66] Greenberg, *Sandworm*, 103.
[67] Greenberg, *Sandworm*, 102-104.
[68] Connell and Vogler, "Russia's Approach to Cyber Warfare," 4.

Despite the absence of formal declarations of cyber war, it is difficult to claim that we exist in a state of cyber peace. Present news cycles are rife with reports of interstate cyber transgressions, ranging from attacks directly attributed to government apparatuses, including the GRU's ongoing password-hacking campaign seeking to access sensitive information held by foreign government and military agencies,[69] to attacks conducted by private hacking organizations which may enjoy tacit government support, such as the 2021 ransomware attack launched by the Russian-speaking group REvil, specifically designed to avoid infecting systems running languages native to the territories of the former USSR.[70] Elsewhere, the US has responded to Russia's digital posturing with cyberattacks of its own, with reports in 2019 claiming that US CYBERCOM had been successful in planting malware within Russia's power grid.[71] Rather than adhering to a norm of cyber non-interference, we have instead entered a sort of cyber-Cold War within which traditional adversaries have traded in conventional tactics for higher tech means of disrupting one another's policies.

Despite its harms manifesting in a physical, and often permanent, capacity, conventional warfare enjoys the advantage of taking place within familiar contexts. The lengthy history of warfare has bestowed the international community with centuries of experience navigating interstate conflicts of all scales, from the grand global campaigns such as both World Wars, to the clandestine proxy conflicts that characterized the Cold War. While the weapons of war have grown more efficient, conventional warfare has remained largely recognizable as direct kinetic conflict between states across the traditional domains of land, air, sea, and, more recently, space. Accordingly, consistent exposure to conventional warfare has led to the establishment of both laws of armed conflict and norms of best practice, both of which are supposed to govern conduct in times of war and peace. International laws pertaining to sovereignty and territorial integrity offer largely clear guidelines for when sovereign states may invoke the right to self-defense.

---

[69] Brian Fung and Zachary Cohen, "Russian Military Targeted Passwords in Wide-Ranging Hacking Campaign, US and UK Officials Say," *CNN*, July 1, 2021, accessed July 8, 2021, https://www.cnn.com/2021/07/01/politics/russian-military-hacking-campaign-us-uk-advisory/index.html.

[70] Ken Dilanian, "Code in Huge Ransomware Attack Written to Avoid Computers That Use Russian, Says New Report," *NBC News*, July 7, 2021, accessed July 8, 2021, https://www.nbcnews.com/politics/national-security/code-huge-ransomware-attack-written-avoid-computers-use-russian-says-n1273222.

[71] David E. Sanger and Nicole Perlroth, "*U.S. Escalates Online Attacks on Russia's Power Grid,*" *The New York Times*, June 15, 2019, accessed July 8, 2021, https://www.nytimes.com/2019/06/15/us/politics/trump-cyber-russia-grid.html.

Further treaties such as the Hague- and Geneva Conventions serve to protect the welfare of combatants and civilians through legally binding articles, while normative considerations such as "probability of success" seek to fill in the gaps that are not explicitly addressed through law. Although transgressions of these laws and norms undoubtedly occur, these measures nevertheless offer a strong and comprehensive governing framework for conduct within conventional conflicts, including distinguishing permissible from impermissible types of attacks, as well identifying when aggression crosses the threshold at which nations may reasonably retaliate in self-defense.

The migration of hostilities from the physical to the digital domain poses particular problems for our understanding of *jus ad bellum* (i.e., "the justice of war" in this classical moral and legal sense of when it is permissible to resort to warfare). The traditional Westphalian notions of sovereignty and territorial integrity, integral to discussions regarding a state's right to self-defense, suffer from concerns of translatability due to the nature of ownership and territory within cyberspace. The digital domain lacks the same clearly demarcated boundaries as the conventional domains of conflict which often stem from internationally recognized land and sea borders, as well as airspace. While ownership of physical equipment such as routers and servers proves uncontroversial given they reside within a state's territory, there is greater controversy surrounding ownership of the intangible elements, such as data, which inhabit cyberspace. Practices such as web hosting, meaning a website originating and belonging to one state being hosted on servers within the territorial confines of another, further render it difficult to discern when a sovereign's digital territory has suffered from an unlawful intrusion. Insofar as the right to defend oneself against an attack is commonly cited as just cause for going to war,[72] the unique realities of territory and sovereignty within cyberspace need to be addressed prior to determining whether cyberattacks may reasonably constitute grounds for going to war.

Furthermore, cyberattacks are themselves often disanalogous to their kinetic counterparts. The conduct of war across the domains of air, land, and sea is readily observable; many of these conflicts are conducted via kinetic weapons and traditional deployments of military force. Likewise, the effects of conventional tactics are immediately apparent as the scale of damage following a conventional attack is straightforwardly observable and quantifiable for purposes of

[72] Brian Orend, "Michael Walzer on Resorting to Force," *Canadian Journal of Political Science*, Vol. 33. Issue 3. (September 2000), 6.

subsequent retaliation, reparation, or declaration of war. This is less the case for cyberattacks. Most cyberattacks are conducted in a clandestine fashion, employing tactics that remain opaque until reverse-engineered by analysts in the days or even months following the attacks. Likewise, the true effects of cyberattacks are often difficult to ascertain, as the task of finding analogues for some of the damages caused by cyberattacks is in many cases impossible. Cyberweapons prove capable of causing immense disruptions across infrastructure, albeit disruptions that can be ceased with a few keystrokes rather than months of physically rebuilding conventionally sabotaged infrastructure. Cyberweapons are also capable of causing forms of harm which are alien to conventional attacks, such as the insidious manipulations of foreign democratic processes. While certain elements of cyberattacks echo elements of their conventional counterparts, it is nonetheless necessary to consider the specific nuances that cyberattacks introduce into the equation. As warfare takes on a new dimension, so too does the decision calculus of conflict.

Firstly, cyberwarfare tends to further blur the already muddied distinction between military and civilian targets. Conventional warfare allows for attacks against legitimate targets. While these are often military in nature, such as air force bases or naval ports, they may also encompass civilian infrastructure deemed critical to an adversary's war effort. For example, civilian rail lines may prove legitimate military targets if they prove integral to the transportation of armour or munitions. Other elements of civilian infrastructure, such as hospitals, present illegitimate military targets--even if targeting them may serve to accelerate an end to hostilities by virtue of dramatically diminishing a state's will to fight. While these distinctions may already prove controversial within conventional warfare, the demarcation grows even more difficult within the realm of cyberwarfare. Military and civilian information networks often intersect, with military cyber operators potentially requiring the usage of civilian network infrastructure to launch cyber offensives. This proves problematic as, insofar as these civilian networks are integral to a nation's cyberwarfare strategy, they may prove to be legitimate military targets for retaliatory efforts by a state seeking to diminish an adversary's cyber capabilities.[73] Although retaliation may intend to simply hamstring a state's ability to wage cyberwar, the targeting of

[73] Patrick Lin, Fritz Allhoff, and Keith Abney, "Is Warfare the Right Frame for the Cyber Debate?", In *The Ethics of Information Warfare*, eds. Luciano Floridi & Mariarosaria Taddeo (Heidelberg: Springer International Publishing, 2014), 42.

civilian network capabilities could potentially result in the shutdown of vital civilian infrastructure, functionally crippling society.

Secondly, cyberwarfare is plagued by the so-called "problem of attribution." The clandestine nature of cyberwarfare can render it difficult to ascertain responsible parties at the time of a cyberattack, particularly if the attack itself is designed to disorient and disrupt an adversary. Whereas conventional attacks are often launched from known military infrastructure such as air bases or naval ships, both of which can be linked to the aggressor directly, cyberattacks can be launched from just about anywhere, requiring little more than a laptop. In certain cases, cyberattacks can even be facilitated using "zombie networks" of thousands of previously infected computers worldwide, rendering it difficult to trace the true origin of an attack both in time and in space. The fog of war surrounding cyberwar is further compounded by the employment of proxies, such as CyberBerkut, conducting cyber operations on behalf of state actors, while preserving some degree of plausible deniability for the benefactors pulling the strings. It is worth nothing that the attribution problem is not absolute as retrospective analysis of cyberattacks often reveals incriminating clues as to the identity of the perpetrators; however, due to the complexity and the time-consuming nature of the discovery process, the attribution problem nonetheless presents an elevated level of risk at the time of the attack. Just as conventional attacks often necessitate a swift defensive response, so too can cyberattacks motivate a prompt retaliatory strike. The attribution problem serves to increase the risk of mistaken retaliation in the immediate wake of a cyberattack, as misdirection may cause a victimized nation to focus its ire against an implicated state, rather than the true belligerent.

Thirdly, the position of cyberwar within the hierarchy of aggression is poorly understood at present. While kinetic operations are typically considered a measure of last resort once all other options have been exhausted, cyber operations have hitherto been seen as a less restrictive option for foreign policy. Espionage, disinformation, and digital sabotage all present measures short-of-war which can be readily employed for the advancement of a nation's interests abroad or to drastically hinder an adversary's ability to do the same. This mindset has underpinned much of the modern military strategies of the traditional superpowers, as evidenced by the growing emphasis on cyber-readiness and the climbing number of recorded cyber incidents. The development of cyberweapons offers states access to flexible political tools which can be more potent than economic sanctions while still falling short of causing the type of sharp, serious

physical damage which would be deemed sufficient for invoking the traditional right to armed self-defense. As the world shifts away from conventional military strength, progressively more interstate conflicts will be shaped or outright won via cyberwarfare.

Fourthly, despite the growing pre-eminence of cyberweapons within interstate conflict, there remains a notable absence of guidelines regarding the permissibility of their use and what responses they would warrant. While many treaties to dictate the permissibility of kinetic weapons have been drafted over the years, such as the 1983 Convention on Certain Conventional Weapons, no such analogues exist to govern cyberweapons. Furthermore, existing treaties drafted specifically to engage with cyber activities, including the Convention on Cybercrime in 2001, fail to encompass interstate cyber conflict within their scope.[74] This legislative void is further exacerbated by the lack of explicit cyber policy on the part of the relevant world cyberpowers. Many nations, such as the United States and the Netherlands, have stated that they would treat severe cyberattacks launched against them as equivalent to armed attacks, leaving open the possibility of responding with kinetic force.[75] However, while these nations have been clear in expressing their willingness to employ kinetic countermeasures, they have proven significantly less forthcoming in terms of defining the thresholds that must be met for a severe cyberattack to transcend being merely a cyber crime and becoming an event of sufficient gravitas so as to merit invoking the right to conventional self-defense. Likewise, very little has been said regarding considerations of proportionality when discussing potential responses. While it is readily imaginable that some potential cyberattacks may prove serious enough to warrant conventional responses, it is less reasonable to assume that conventional responses would be proportional in all cases.

The ambiguity of state tolerance of cyberwarfare extends beyond policy and into practice. While serious cyberattacks have been theorized to be equivalent to armed attacks, there nonetheless exists a range of cyber hostilities which have failed to garner strong reactions amongst the international community. Acts of cyberespionage are largely tolerated, having been widely regarded as measures well short of war. Cyber disinformation campaigns have likewise

[74] Michael N. Schmitt and Liis Vihul, "The Emergence of International Legal Norms for Cyberconflict," in *Binary Bullets: The Ethics of Cyberwarfare*, eds. Fritz Allhoff, Adam Henschke, & Bradley Jay Strawser (New York, NY: Oxford University Press, 2016), 38-39.
[75] Schmitt and Vihul, "The Emergence of International Legal Norms for Cyberconflict," 42.

been met with little more than stern admonition, suggesting these attacks similarly fall short of crossing the undefined threshold. Even the international community's approach to cyber sabotage has proven to be quite nebulous. Amidst the sweeping DDoS attacks that shut down much of its digital infrastructure in 2007, Estonia argued that the attacks were sufficient for invoking Article 5 of the North Atlantic Treaty, suggesting that the cyberoperation was regarded as potentially constituting an armed attack; NATO would disagree with the Estonian assessment, suggesting the attacks failed to cross the threshold which would warrant invoking the right to collective defense.[76] Evidently, the temporary paralysis of a nation's infrastructure proves insufficient. Despite resulting in physical destruction, the Stuxnet attack of 2010 likewise failed to inspire regulatory discourse within the international community. Instead, the global cyberpowers chose to remain conspicuously quiet regarding Stuxnet. Even Iran opted not to raise the issue of Stuxnet to the UN Security Council, despite having been targeted by the cyberweapon. Dipert notes that the international response regarding Stuxnet, or the lack thereof, suggests a tacit international agreement that Stuxnet-esque attacks may constitute the upper end of the spectrum of tolerable cyberattacks.[77] Accordingly, even acts of cyber sabotage which cross the digital Rubicon and result in tangible destruction have hitherto been seen as not necessarily acts of war. The potential for governing cyberweapons is further hamstrung by a general sense of reluctance amongst cyberpowers to proactively restrict *their own* cyberweapons--which may yet prove to give them a significant advantage on the battlefield and perhaps elsewhere.[78] Evidently, the international community's present policies on cyberwar are amorphous at best.

In the absence of a governing framework for cyberwar conduct, we find ourselves ill-equipped to negotiate the landscape of virtual hostilities. Without consensus regarding the acceptable thresholds of cyber harm, it becomes difficult to discern not only when just cause for retaliation has been met, but also what permissible form such retaliation may take. While kinetic attacks may reasonably beget kinetic responses, it remains to be seen whether the threat of conventional retaliation promised by some cyberpowers may ever be warranted in the face of a cyberattack. Given the present lack of impetus amongst relevant stakeholders to commit to a governing framework, Dipert argues that we are currently headed towards a state of 'game-

[76] Lucas, "Emerging Norms for Cyber Warfare," 26.
[77] Randall R. Dipert, "Distinctive Ethical Issues of Cyberwarfare," In *Binary Bullets: The Ethics of Cyberwarfare*, eds. Fritz Allhoff, Adam Henschke, & Bradley Jay Strawser (New York, NY: Oxford University Press, 2016), 64.
[78] Schmitt and Vihul, "The Emergence of International Legal Norms for Cyberconflict," 40.

theoretic equilibrium' with regards to cyberwarfare.[79] The power asymmetry between cyber offense and defense, coupled with the increased costs of a purely defensive cyber posture, ensures that cyberpowers will prioritize the development of cyberoffensive capabilities. Meanwhile, the contemporary attitude of the global community of cyberpowers suggests general reluctance to proactively restrict cyber capabilities. What results is a digital grey zone within which cyberpowers are free to continue developing and deploying their offensive cyberweapons, probing their adversaries to see where their threshold of tolerance lies. In the absence of an overarching governing framework, Dipert suggests that the cyber frontier may resemble a "cold war" within which nations shall tend to avoid using their most destructive weapons, while a semblance of stability remains in place due to the promise of Mutual Assured Destruction (MAD) made possible through cyber means.[80]

2.4 Beyond Dipert's Equilibrium

While Dipert suggests that this quite indirect norm of avoiding the worst death and widespread destruction seems to be developing in the field of cyberwarfare,[81] I argue that this degree of stability alone is insufficient. While cyber operatives have hitherto avoided causing tangible destruction and deaths, the current status quo of probing an adversary's tolerance runs a real risk of mistakenly crossing a poorly defined threshold for self-defense and causing conventional war to break out; while most states seem to mark a threshold for self-defense at being subject to substantial physical damage and/or deaths, it remains to be seen whether *each* nation sets their threshold for self-defense that high. Furthermore, cyberwar stands apart from its conventional kin by virtue of being readily able to inflict numerous and ongoing harms *beyond* immediate casualties. As evidenced in Estonia and Ukraine, cyberattacks are more than capable of shutting down entire state infrastructures, rendering it impossible for civilians within them to access basic necessities such as proper medical care and electricity. The 2021 attacks in the US have similarly revealed serious vulnerabilities in the supply of oil and gasoline, as well as that of clean drinking water. Cyberattacks likewise can be, and indeed have been, used to sabotage

[79] Dipert, "Distinctive Ethical Issues of Cyberwarfare," 70.
[80] Dipert, "Distinctive Ethical Issues of Cyberwarfare," 70.
[81] Dipert, "Distinctive Ethical Issues of Cyberwarfare," 70.

democratic processes through attacks on election infrastructure and the wholesale manipulation of information. While neither result in physical damage, it is difficult to argue that neither has the potential of crossing the threshold of an unacceptable attack on a sovereign state. There are also concerns regarding the lack of targeting restrictions for cyberweapons as cyberattacks have hitherto shown little concern for distinguishing between military, government, and civilian targets; cyber weapons such as NotPetya have proven largely indiscriminate, designed to infect as many machines as possible and significantly increasing the risk of unforeseen collateral harms on defenseless civilian populations. As a result, merely leaving things, in laissez-faire fashion, to tend towards a game-theoretic equilibrium within the cyber domain carries with it undue risk, necessitating a governing framework to mitigate the likelihood of potentially disastrous consequences. The next chapter will consider in detail one such recent attempt—with high profile and expert authority—to craft such a cyber-governing framework. This attempt would take the form of the *Tallinn Manual*.

Chapter 3

Extending Law into Cyberspace: The *Tallinn Manual*

3.1 Escaping a Cyber State of Nature

While the threat of cyberattacks grew steadily in public consciousness throughout the early 2000s, comparatively little would be done to address the digital arms race within the cyber domain until 2008. Although the effects of the 2007 DDoS attacks on Estonia would ultimately prove relatively mild once the dust had settled, the unprecedented widespread disruptions of Estonia's cyber infrastructure served to announce cyberwar as a clear and imminent threat not only for traditionally vulnerable states, but also the hitherto untouchable Western nations of NATO. Furthermore, the ensuing disagreement between Estonian authorities and NATO officials regarding the appropriate designation for the attacks rendered it starkly apparent that the (then-obtaining) status quo of cyber governance was untenable in the long term. The absence of a governing framework, be it legal or ethical, ensured that achieving evaluative consensus regarding cyberattacks would remain a borderline impossible task. As evidenced by the Estonia attacks, the immediately crippling effects of a severe cyber-strike might motivate a victimized state to designate it, for all intents and purposes, an armed attack against its core national interests. Concurrently, an international body such as NATO, removed by degrees from the immediate epicenter of the attack, might regard the same cyber operation as merely an offense warranting a diplomatic response, still falling well short of the threshold for justifiably invoking the right to self-defense.

These differences in perception only serve to amplify the effectiveness of offensive cyber operations as targeted states face challenges on two fronts. First, the general difficulties of establishing effective cyber defenses, coupled with the comparative potency of cyberweapons, all but ensures that aggressive cyberoperations will succeed in achieving at least some of their objectives. Second, the general ambiguity surrounding the permissibility of cyberattacks and their position within the hierarchy of aggressive action limits a state's ability to respond in a timely and effective manner. This leaves the victim state vulnerable to continuing harms by an aggressor as lengthy investigations and dialogues seek to discern whether the incoming

belligerent cyber operation constitutes a use of force or armed attack. The immediacy of cyberattacks presents a stark contrast to the often-lengthy nature of diplomatic processes. The lethargic response time of diplomatic efforts is itself further aggravated by the nebulousness of distinctions, perceptions, and thresholds regarding cyberattacks, ensuring difficulty in achieving international consensus in an expedient manner. These political complications serve to further increase the effectiveness of cyberattacks as potential responses are all but guaranteed to be delayed, potentially well after the cyber operation has achieved its objectives and has even been discontinued by the aggressor state.

The general turmoil surrounding the international response (or lack thereof) to the Estonian attacks motivated the 2008 establishment of the NATO Cooperative Cyber Defence Centre of Excellence (CCD COE) in the Estonian capital of Tallinn. Staffed by an assortment of military and civilian advisors, the CCD COE was tasked with cyberwar research and analysis to both shape NATO's cyber doctrine, as well as to improve the organization's general defensive readiness and ability to respond to belligerent cyberattacks against its member states. The CCD COE wasted little time launching an initiative with the express intent of addressing the vacuum of cyberwar governance which had exacerbated disagreement between Estonia and NATO in 2007, to the embarrassment of both and to the quiet delight of Russia, who no doubt took special note of the chaotic Western non-response. Establishing some form of governing framework would, in theory, provide clearer guidelines regarding the differentiation of permissible and impermissible cyberattacks, as well as demarcate the threshold at which a nation may retaliate in the interests of self-defense. To this end, the CCD COE invited a team of prominent independent experts in the field of international law (hereafter referred to as simply 'the Experts') to undertake the drafting of a manual outlining the applicability of existing international law within the nascent domain of cyberwar.[82]

The conclusions drawn by the Experts would ultimately take the form of the *Tallinn Manual on the International Law Applicable to Cyberwarfare*, first published in 2013 and subsequently revised and rereleased as *Tallinn 2.0* in 2017 to encompass cyber crimes beyond the scope of interstate conflict. Despite being drafted by a team of international legal experts, the

[82] Michael N. Schmitt and NATO Cooperative Cyber Defence Centre of Excellence, eds. *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations*. Second edition. Cambridge, United Kingdom; New York, NY, USA: Cambridge University Press, 2017. 3.

Tallinn Manual is itself not a binding legal document introducing new cyberwar legislation. Rather, the intent of the *Tallinn Manual* is to provide close legal analyses of existing laws of armed conflict (LOAC) to gauge how, if at all, they may apply to cyberwarfare. Beginning from a position of consensus regarding the applicability of international laws to the cyber domain, the Experts would structure the *Tallinn Manual* along a series of black-letter rules for cyber conduct, each derived from presently binding international law and legal precedents.[83] Each rule would be formulated based on Expert consensus and coupled with brief commentary offering insight into the Experts' rationale for its adoption, or construction, as well as any potential disagreements that had arisen during the deliberation process. The rules themselves were presented in a manner reminiscent to that of treaties, as the Experts sought to interpret existing law in a readily recognizable manner. While the *Tallinn Manual* itself holds no legal force, the Experts hoped to show that, insofar as the rules contained within the *Tallinn Manual* accurately reflected already-binding laws of armed conflict, they could and should be seen as being "binding on all States" in connection with cyber-conflict.[84]

While the scope of *Tallinn 2.0* encompasses both *jus ad bellum* and *jus in bello*, as well as a myriad of peacetime cyber operations which fall beyond the considerations of either framework (such as cyberattacks by non-state entities), the following sections will focus on the document's work pertaining to *jus ad bellum*. In particular, I will focus on *Tallinn 2.0*'s rules regarding sovereignty and the use of force within cyberspace due to the pivotal role these considerations play in the United Nations' (UN) own *jus ad bellum* conclusions. The principle of sovereignty and the protections it affords present an element integral to the UN's deliberations regarding potential violations of Article 2(4) of the UN Charter prohibiting unjustified interference into the affairs of sovereign states; the UN's definition of aggression is likewise derived from the assertions made within Article 2(4).[85] Accordingly, the early rules in *Tallinn 2.0* seek to lay the groundwork for how these conceptions of sovereignty translate into cyberspace and how cyber operations might feasibly be considered violations of Article 2(4).

---

[83] *Tallinn Manual 2.0*, 3-4.
[84] *Tallinn Manual 2.0*, 4.
[85] United Nations, *Charter of the United Nations*, October 24, 1945, 1 UNTS XVI.
UN Charter, Article 2(4): "All Members shall refrain in their international relations from the threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the Purposes of the United Nations."

Tallinn 2.0's later rules regarding use of force work towards identifying which kinds of cyber operations may constitute uses of force, while further contemplating what threshold of severity must obtain before a digital operation could motivate immediate acts of self-defense as permitted, importantly, by Article 51.[86] These two sets of rules, taken in tandem, serve to address each of the four principles of *jus ad bellum* specified within the LOAC: namely, just cause, proportionality, last resort, and public declaration by proper authority.[87] What results is a strong evaluative framework which works towards redressing the lack of governance within the cyber domain. The aim of this chapter is to both explain and evaluate the contents of *Tallinn 2.0* in this regard.

3.2 Sovereignty and Territorial Integrity in Cyberspace

From its outset, *Tallinn 2.0* stringently expresses the importance of upholding state sovereignty even with the transition to cyberspace, with the Experts positing sovereignty as the "foundational principle of international law".[88] The Experts note that international laws concerned with matters of domestic jurisdictions, the prohibition of force, and global non-interventionism, all derive from this overarching principle of sovereignty.[89] This sentiment is echoed in the early sections of the UN Charter, with Article 2(1) asserting that each of the UN's member states enjoys the right to sovereignty in equal measure with each of its peers.[90] Any semblance of legal stability within the international community is predicated on the unanimous acknowledgement of states as sovereign entities, each entitled to the same set of inherent rights and bound by the same collection of duties to one another. Without the principle of sovereignty, international law would prove largely untenable.

[86] UN Charter, Article 51: "Nothing in the present Charter shall impair the inherent right of individual or collective self-defence if an armed attack occurs against a Member of the United Nations, until the Security Council has taken measures necessary to maintain international peace and security. Measures taken by Members in the exercise of this right of self-defence shall be immediately reported to the Security Council and shall not in any way affect the authority and responsibility of the Security Council under the present Charter to take at any time such action as it deems necessary in order to maintain or restore international peace and security."
[87] Orend, *War and Political Theory*. (Cambridge, UK; Medford, MA: Polity, 2019), 82.
[88] *Tallinn 2.0*, 11.
[89] *Tallinn 2.0*, 11-12.
[90] UN Charter, Article 2(1): "The Organization is based on the principle of the sovereign equality of all its Members."

Accordingly, *Tallinn 2.0*'s first rule is an assertion that sovereignty exists within the cyber domain, rather than being solely a feature of real space.[91] While this represents a seemingly uncontroversial proposition, this claim would nonetheless trigger the need for further justification on the part of the Experts. Our traditional Westphalian conceptions of sovereignty have hitherto been intricately tied to territorial integrity; the Experts themselves were quick to offer a definition of sovereignty guaranteeing independence, allowing nations to serve the functions of a state within the confines of a specified territory, free of external interference.[92] Accordingly, unlawful intrusions by one state into the territory of another, whether military in nature or otherwise, are regarded as violations of the latter's sovereignty on these grounds. While still occasionally proving controversial, the distinction of when a violation of sovereignty has occurred is made more clear-cut by the existence of explicitly demarcated borders marking where one state's territory ends and another's begins; the movement of armed troops, for example, constitutes a violation of traditional sovereignty if these national borders are unlawfully breached. However, this task quite literally takes on a new dimension within the cyber domain as concepts such as territory and ownership become much harder to parse than in real, physical space. Does "territory" within cyberspace exist in a Westphalian sense? What corners of cyberspace, if any, can a state legitimately claim to own, and on what grounds? How is jurisdiction measured when considering 'domestic' sites hosted on physical servers within the borders of another sovereign state? It is these kinds of questions which pose concerns and perplexities for *Tallinn 2.0*'s prime rule and necessitate an array of accompanying Expert comments. It is through this commentary that the Experts extend our Westphalian notions of sovereignty into cyberspace and assert that Rule 1 indeed applies (and, in turn, motivates the subsequent rules which derive from it).

The Experts sought to assuage concerns of sovereignty's translatability into cyberspace by arguing that cyber sovereignty exists across three specific layers: the physical, the logical, and the social.[93] The physical layer of cyber sovereignty is most readily recognizable through the lens of Westphalian territoriality. It consists of the physical elements which enable network technology, including devices such as routers, servers, computers, and the relevant adjacent

[91] *Tallinn 2.0*, Rule 1: "The principle of State sovereignty applies in cyberspace."
[92] *Tallinn 2.0*, 11.
[93] *Tallinn 2.0*, 12.

infrastructure.[94] Insofar as these kinds of devices are physically located within a sovereign's territory, they fall under the purview of the traditional protections afforded by territorial sovereignty. In contrast, the logical layer refers to the data, applications, and other entities which exist within cyberspace.[95] Accordingly, this refers to the intangible elements residing within the cyber domain itself, such as websites and databases, being made possible by hardware in the physical layer. Finally, the social layer refers to the users of the digital domain, be they individuals or groups.[96] This layer includes citizens engaging with cyberspace in some capacity, as well as both government and non-government organizations conducting their business in the cyber domain. Much as a sovereign enjoys authority over its citizens and territories in real space, so too does it possess the same jurisdictional authority in cyberspace across each of these three layers.

Having reasserted that the principle of sovereignty continues to apply within cyberspace, the Experts drafted Rules 2, 3, and 4 to follow Rule 1 to its logical conclusions and establish a cyber norm of non-interference, much like that which governs international law within real space. Rules 2 and 3 divide sovereignty into internal and external elements, respectively. Rule 2 works towards ensuring that states enjoy the same independence to govern domestically that they are afforded by the principle of sovereignty in the physical domain.[97] Accordingly, internal sovereignty enables sovereigns to fulfill the traditional functions of a state as they pertain to cyberspace. This includes freedom to implement regulatory measures over cyber activities and actors deemed to be domestic (however, the Experts note that this obtains insofar as such measures are compliant with other binding international laws, such as human rights laws).[98] In addition to this regulatory role, internal sovereignty likewise grants states the right to defend their domestic cyber interests as they may manifest across the three layers of cyber sovereignty.[99] With regards to the accompanying external sovereignty in the cyber domain, Rule 3 states that sovereign nations are free to conduct international cyber activities, provided they do not infringe

[94] *Tallinn 2.0*, 12.
[95] *Tallinn 2.0*, 12.
[96] *Tallinn 2.0*, 12.
[97] *Tallinn 2.0*, Rule 2: "A State enjoys sovereign authority with regard to the cyber infrastructure, persons, and cyber activities located within its territory, subject to its international legal obligations."
[98] *Tallinn 2.0*, 13.
[99] *Tallinn 2.0*, 13.

on any other existing international laws.[100] This rule grants states the ability to act within the larger cyber domain as they see fit, including those activities which extend beyond the immediate confines of their own territory, albeit with some caveats.[101] While external sovereignty affords states a great degree of liberty in terms of what cyber activities are made permissible for them, limitations arise when one state's external sovereignty clashes with another's internal sovereignty.

Declaring that internal sovereignty takes precedence over external sovereignty in such cases, Rule 4 of *Tallinn 2.0* declares that states are not permitted to conduct cyber operations which would serve to violate the internal sovereignty of another state.[102] In keeping with the conception of sovereignty as a state's right to independence and to fulfill the function of a state within its own territory, the Experts argue that cyber operations impeding a state's ability to fulfill this function constitute violations of sovereignty.[103] Echoing existing laws prohibiting state interference in the affairs of sovereign states (barring exceptions such as UN sanctioned interventions or wars of self-defense), Rule 4 reasserts that states possess the same duty to respect one another's sovereignty within the digital domain. Accordingly, Rule 4 prohibits state cyber operations which prove violative of any of the three layers of another state's cyber sovereignty. This rule likewise brings *Tallinn 2.0* firmly in line with the UN Charter's Article 2(4). Much like Rule 4, Article 2(4) of the UN Charter moves to establish a rule of non-interventionism amongst states.[104] By defending the position that the principle of sovereignty underpinning international law remains applicable within the cyber domain, the first four rules of *Tallinn 2.0* establish that states are bound by the same principles of non-interference which govern conduct within the conventional domains of conflict.

Provided that they were the efforts of a state or state-backed entity, cyber operations such as the DDoS attacks on Estonia and the NotPetya worm, which effectively shut down the Ukrainian government, may be regarded as violations of Rule 4. In both cases, the cyberattacks proved violative of the internal layers of a state's sovereignty. The attacks on Estonia's

[100] *Tallinn 2.0*, Rule 3: "A State is free to conduct cyber activities in its international relations, subject to any contrary rule of international law binding on it."
[101] *Tallinn 2.0*, 16.
[102] *Tallinn 2.0*, Rule 4: "A State must not conduct cyber operations that violate the sovereignty of another State."
[103] *Tallinn 2.0*, 17.
[104] UN Charter, Article 2(4).

infrastructure served to shut down a wide range of government sites, while simultaneously hitting native Estonian users and businesses occupying the social layer of Estonia's internal sovereignty. Likewise, the attacks on Ukraine served to paralyze its domestic government, with the NotPetya worm actively working to destroy data falling under the purview of the logical layer of the state's internal sovereignty.[105] Both cyberattacks further served to hinder the ability of the state to efficiently fulfill its government functions. Given the growing digitization of healthcare and welfare distribution processes, such processes now find themselves susceptible to being significantly hindered or halted entirely by cyber offensives; for example, the WannaCry attacks described last chapter serve to offer a brief glimpse into how foreign cyber actors can directly influence a state's ability to distribute healthcare to its citizens. While none of these attacks may be seen as immediately justifying the right to armed self defense, each undeniably illustrates how cyber operations may potentially violate the digital sovereignty of states.

Much like Article 2 of the UN Charter, the first four rules of *Tallinn 2.0* also serve to open the door to the possibility that cyber operations may in fact constitute acts of aggression. The UN's own definition of aggression, which in turn motivates its deliberations regarding conflict arbitration and international peacekeeping, is built upon the principles originally outlined in Article 2. In 1974, the UN General Assembly would adopt Resolution 3314, Article 1 of which defines aggression as "the use of armed force by a State against the sovereignty, territorial integrity, or political independence of another state or in any other manner inconsistent with the Charter of the United Nations, as set out in this Definition".[106] Notably, Article 2 of the resolution goes on to state that some such acts may nonetheless be regarded as acts short of aggression should the Security Council decide that the associated consequences are of an insufficient scale.[107] While Resolution 3314 offers some classic examples of acts of aggression for illustrative purposes, such as kinetic attacks against a state's territory or blockades of their ports, the resolution acknowledges that such a list is not exhaustive and that other acts may well

[105] Andrew E. Kramer, "Ukraine Cyberattack Was Meant to Paralyze, Not Profit", *The New York Times*. June 28, 2017, Accessed September 20, 2021, https://www.nytimes.com/2017/06/28/world/europe/ukraine-ransomware-cyberbomb-accountants-russia.html.
[106] UN General Assembly, *Definition of Aggression*, 14 December 1974, A/RES/3314.
To clarify: UN General Assembly Resolutions are not normally considered binding international law. Yet they can have moral suasion, and the import of this often-cited Resolution 3314 is that it was passed to further clarify the crucial concept of aggression, understood as a violation of state sovereignty so severe that a forceful armed response may justly be undertaken.
[107] UN General Assembly Resolution 3314.

constitute further acts of aggression.[108] As evidenced in the preceding chapter, the advent of cyberwarfare has rendered it possible to achieve similar deleterious effects to those achieved by conventional means. Whereas states may have once relied on military blockades of ports, they now prove capable of shutting them down remotely through disabling navigation networks, traffic control towers, or other adjacent infrastructure; one such cyberattack on South Africa's state-owned freight company managed to completely shut down their computer systems, forcing them to revert to manually processing shipments and resulting in a significant export backlog.[109] As a result, it would appear possible, if not plausible, that purely digital attacks could qualify as acts of aggression in the eyes of the UN should they 1) violate a state's sovereignty and 2) be of sufficient consequence.

The continued importance of state sovereignty within cyberspace asserted by *Tallinn 2.0*'s rules, as well as the likelihood that violations of cyber sovereignty may be treated as acts of aggression in accordance with the UN's definition, offers an avenue for just cause within the cyber domain. In turn, the possibility of cyber aggression offers the UN the right to potentially intervene in accordance with articles of defense and intervention outlined in Chapter VII of the UN Charter. The provisions within articles 39-51 bestow the UN with the authority to deploy a myriad of measures—diplomatic, sanctions, etc.—to de-escalate a state's aggression peacefully, while maintaining the right to deploy armed force if such measures are exhausted.[110] As a result, the early efforts of *Tallinn 2.0* establish one way within which cyberattacks may meet the just cause condition for armed conflict insofar as the UN may feasibly respond multilaterally to certain types of cyberattacks with armed force should such attacks prove to be grievous violations of state sovereignty. Later work on *jus ad bellum* within *Tallinn 2.0* takes on the difficult task of determining whether the second type of grounds for just cause for war, that of a state's inherent right to self-defense in the face of an armed attack, may conceivably be met by the deployment of cyber weapons.

[108] UN General Assembly Resolution 3314.
[109] Felix Njini, "Crippled Africa Port Goes Manual as PE-Backed Software is Shut", *Bloomberg*. July 22, 2021. Accessed September 24, 2021. https://www.bloomberg.com/news/articles/2021-07-22/south-africa-s-transnet-reports-disruption-to-it-services.
[110] UN Charter, Articles 39-51.

3.3 When Push Becomes Shove: Force and Armed Attacks

Much of the UN's decision calculus relating to *jus ad bellum* revolves around crucial distinctions between *uses of force* and *armed attacks*. A belligerent nation targeting another state with a cyber operation may be judged by the UN as committing a grievous violation of the target's sovereignty, depending on the severity of the associated consequences. This may, in turn, fulfill the just cause requirement for the UN's deployment of armed force in accordance with Article 42, provided the non-armed alternatives outlined in Article 41 have proven insufficient for restoring peace.[111] While the weaker designation of *use of force* may motivate multilateral armed intervention by the UN, the more demanding label of *armed attack* plays the key role in determining whether a state may be justified in unilaterally employing armed force in the interests of self-defense, as outlined within Article 51. Should international law remain binding irrespective of whether the domain is physical or digital, it would stand to reason that these distinctions and the responses they motivate would likewise obtain in cyberspace. Accordingly, in seeking to extend existing international law into cyberspace, the *Tallinn 2.0* Experts next considered conceptualizing how the UN's deliberations regarding use of force and armed attacks translate into the cyber domain.

Much like the initial propositions insisting that sovereignty remains applicable within cyberspace leading to Rules 1-4, an assertion that uses of force may occur in the cyber domain necessarily precedes the associated rules suggested by the Experts. While the employment of conventional weapons is readily seen as constituting at least a use of force, it is less readily apparent whether the usage of cyberweapons could conceivably cross the same threshold. In response to potential skepticism, the Experts adopted a precedent within the 1996 International Court of Justice (ICJ) advisory opinion regarding the legality of nuclear arsenals. While determining whether articles within the UN Charter prohibited the development and acquisition of nuclear weaponry, the ICJ determined that the provisions within the Charter "apply to any use

[111] 1: UN Charter, Article 41: "The Security Council may decide what measures not involving the use of armed force are to be employed to give effect to its decisions, and it may call upon the Members of the United Nations to apply such measures. These may include complete or partial interruption of economic relations and of rail, sea, air, postal, telegraphic, radio, and other means of communication, and the severance of diplomatic relations."
2: UN Charter, Article 42: "Should the Security Council consider that measures provided for in Article 41 would be inadequate or have proved to be inadequate, it may take such action by air, sea, or land forces as may be necessary to maintain or restore international peace and security. Such action may include demonstrations, blockade, and other operations by air, sea, or land forces of Members of the United Nations."

of force, regardless of the weapons employed".[112] Accordingly, the Experts cited this determination as grounds for holding that the digital nature of cyberweapons does not immediately preclude them from being potentially regarded as uses of force akin to conventional weapons.[113] This declaration would, in turn, render it theoretically possible for cyberattacks to serve as just cause for war in accordance with Articles 2(4) and 42 in the case of UN intervention, as well as immediate unilateral self-defense as per Article 51 should the use of force be significant enough to qualify as an armed attack.[114]

With the groundwork laid for regarding cyberweapons as feasibly constituting uses of force and armed attacks, the Experts then moved on to introduce Rules 68-75 to expand on the nuances of force and self-defense in cyberspace. Rules 68-70 primarily concern the prohibition of both the use and threat of force within cyberspace, while Rules 71-75 consider the more severe cases within which a use of force crosses the threshold of being an armed attack and the associated right of self-defense. Once more extending the sentiments of the UN Charter into cyberspace, Rule 68 asserts that cyber operations constituting a threat or use of force, much like their real space counterparts, are likewise prohibited by law. [115] This rule serves to expand the scope of the UN's prohibition on threats and use of force to encompass the threat and use of cyber weapons, or at least those promising a significant quantum of damage and magnitude of consequence. Accordingly, the rule asserts that conventional and cyber uses of force fall under the purview of the same international laws, rather than suggesting they are fundamentally different phenomena.

With the prohibition on force extended into cyberspace, Rule 69 tackles the critical task of defining "use of force" in the cyber context. This would prove a deceptively complex task as the UN Charter itself fails explicitly to define the term, forcing the Experts to look elsewhere for legal precedents from which to draw their own conclusions.[116] Early guidance would once more emerge from past UN proceedings, this time in the form of the 1945 UN Charter drafting conference as well as the 1970 General Assembly's Declaration on Friendly Relations. In

[112] *Tallinn 2.0*, 328.
[113] *Tallinn 2.0*, 328.
[114] *Tallinn 2.0*, 328.
[115] *Tallinn 2.0,* Rule 68: "A cyber operation that constitutes a threat or use of force against the territorial integrity or political independence of any State, or that is in any other manner inconsistent with the purposes of the United Nations, is unlawful."
[116] *Tallinn 2.0,* 330.

deliberations regarding potential uses of force, both proceedings were noted by the Experts as having declined to recognize economic and political coercion as recognizable acts of force.[117] Drawing on this precedent, the Experts ultimately concluded that "neither non-destructive cyber psychological operations intended solely to undermine confidence in a government, nor a State's prohibition of e-commerce with another State designed to cause negative economic consequences, qualify as uses of force".[118] Insofar as conventional coercion tactics along these lines were seen as falling short of the elusive use of force threshold, so too would their cyber equivalents.

However, the opposite would also hold true. The Experts' search for a definition of use of force ultimately turned towards the ICJ's judgement in the 1986 case of *The Republic of Nicaragua v. The United States of America*. Following the US' attempts to arm and train Nicaraguan Contras with the express intent of overthrowing the socialist Sandinista government which had emerged in the wake of the Nicaraguan Revolution, the Nicaraguan government argued that the US' conduct qualified as a use of force and a grievous violation of its sovereignty.[119] In the ensuing legal deliberations regarding whether US involvement constituted an armed attack against the Nicaraguan government, the ICJ stated that the 'scale and effects' of actions needed, importantly, to be taken into account to gauge adequately whether an armed attack had transpired.[120] Although the ICJ referred to scale and effects with specific regards to the armed attack distinction, the Experts reasoned that the same metric may be employed to identify uses of force; insofar as a cyber operation has similar scale and effects to a non-cyber use of force, so too would the cyber operation count as a use of force.[121] The Experts further borrowed from the ICJ's judgement in the *Nicaragua* case and agreed that armed attacks represented "the 'most grave' forms of the 'use of force'".[122] As a result, while all armed attacks are necessarily uses of force, not all uses of force achieve the scale and effects required to be classified as an armed attack.

---

[117] *Tallinn 2.0*, 331.
[118] *Tallinn 2.0*, 331.
[119] *Case Concerning Military and Paramilitary Activities In and Against Nicaragua (Nicaragua v. United States of America); Merits*, International Court of Justice (ICJ), 27 June 1986.
[120] *Tallinn 2.0*, 330.
[121] *Tallinn 2.0.* 331.
[122] *Tallinn 2.0*, 331.

The conclusions garnered from the *Nicaragua* judgement ultimately underpin the Experts' attempts at establishing an evaluative framework for cyber uses of force. Pointing towards both thresholds of harm and specific qualitative elements of cyberattacks, the Experts adopt an analogy approach towards distinguishing cyber uses of force from less-than-forceful cyber operations. Under this approach, cyberattacks resulting in harms similar to those traditionally wrought by conventional attacks, namely bodily harm and physical destruction, are uncontroversially regarded as uses of force. In the event of cyber operations failing to cause physical harms, the Experts suggest a series of qualitative criteria which may be taken into account to determine whether a given cyber operation is likely to be viewed by the international community as a use of force.[123] Among the non-exhaustive criteria offered by the Experts are consequentialist considerations of the severity of the operation's consequences, the immediacy with which the consequences arise, as well as the invasiveness of the operation.[124] In the absence of physical harm, this array of criteria can be used to determine whether a cyber operation is analogous in its scale and effects to the kinds of conventional operations which would be widely regarded as uses of force. Ultimately, the force designation of a cyber operation is derived from the precedent set by analogous conventional operations, rendering it uncontroversial to distinguish a cyber operation as a use of force should its consequences prove similar to conventional alternatives.

Following a brief supplementary definition of threat of force in Rule 70,[125] Rules 71-75 seek to address self-defense within the cyber domain. Rule 71 works towards extending the self-defense provisions laid out in Article 51 of the UN Charter into cyberspace, maintaining a state's inherent right to self-defense in the event of cyberattacks that have risen to the level of armed attacks.[126] As was the case with uses of force, the Experts would assert that the distinction of armed attack is not exclusive to operations of a kinetic nature; rather, they readily acknowledge that purely cyber operations likewise have the potential of rising to the level of an armed attack

[123] *Tallinn 2.0*, 333.
[124] *Tallinn 2.0*, 333-335.
[125] *Tallinn 2.0*, Rule 70: "A cyber operation, or threatened cyber operation, constitutes an unlawful threat of force when the threatened action, if carried out, would be an unlawful use of force."
[126] *Tallinn 2.0*, Rule 71: "A State that is the target of a cyber operation that rises to the level of an armed attack may exercise its inherent right of self-defence. Whether a cyber operation constitutes an armed attack depends on its scale and effects."

sufficient for the purposes of Article 51.[127] Once more, this distinction would be contingent on a consequentialist evaluation of a cyber operation's effects. In the view of the Experts, provided the scale and effects of a cyberattack "were analogous to those that would result from an action otherwise qualifying as a kinetic armed attack", the cyberattack would itself qualify as an armed attack.[128]

This leaves open the question of what threshold must be crossed before the scale and effects of a cyber operation are of a sufficient magnitude to be considered an armed attack. Acknowledging that codified law lacks a formal set of criteria for this distinction, the Experts once more employ analogues to demonstrate which sort of cyber operations they view as uncontroversially constituting armed attacks and which types of operations fall short of the distinction. To this effect, operations geared towards cyber espionage are seen as measures short of armed attacks. Likewise, disruptive attacks targeting non-essential cyber services are similarly deemed to fail to cross the threshold of being armed attacks; the DDoS attacks on Estonia's private cyber infrastructure which caused disruptions to local businesses and community forums fall short of being "armed" on these grounds. Conversely, the Experts arrived at consensus agreement that cyber operations resulting in significant physical harm and/or destruction are recognizably armed attacks, irrespective of their digital origin.[129] In the absence of formal legislation demarcating when a use of force becomes an armed attack, the Experts adopt a more normative and open-ended approach to evaluating the range of cyberattacks; those types of operations which have hitherto been treated by the international community as falling short of armed attacks in real space should continue to be treated as such when they occur in cyberspace. In contrast, cyber operations with the same scale and effects of kinetic armed attacks are themselves treated as armed attacks sufficient even for invoking the inherent right of self-defense outlined in Article 51. Accordingly, a devastating cyberattack resulting in the remote detonation of a state's weapon stockpiles would be evaluated as being essentially equivalent to a kinetic strike on the same target, potentially warranting an Article 51 response.

Although Rule 71 defends the right to self-defense in the face of armed cyberattacks, the Experts note that there are subsequent requirements of necessity, proportionality, imminence,

[127] *Tallinn 2.0*, 340.
[128] *Tallinn 2.0*, 341.
[129] *Tallinn 2.0*, 341.

and immediacy which must be met prior to responding with force in a justified manner.[130] These requirements are expanded on within Rules 72 and 73, with the former asserting that any forceful retaliatory response brought to bear by a victimized state must be both necessary and proportionate.[131] The first requirement of necessity enables a state to respond with force if and only if the use of force is necessary to stop an armed attack against the state. This requirement does not preclude the responsibility of victim states to explore non-forceful alternatives first. Should an incoming cyberattack be potentially mitigated by short-of-force operations, such as the redoubling of cyber defenses or the threat of severe sanctions, the victim state would not be justified in deploying force. Only once non-forceful options have been exhausted can a state respond to an armed attack with force of its own. Notably, the Experts do not differentiate between the permissibility of kinetic- and cyber uses of force *once the necessity condition has been met*; both kinetic and cyber responses are considered potentially acceptable responses.[132] Meanwhile, the principle of proportionality tempers prospective uses of force once the necessity condition has been met. The principle of proportionality demands that the retaliatory use of force remains at an acceptable level relative to the threat. The force of a victim's response needs not be exactly commensurate with the force of an aggressor's armed attack, as the Experts note that a greater or lesser degree of force may prove necessary to deter an aggressor's efforts.[133] However, the proportionality requirement is designed to ensure that the ensuing response, be it kinetic or cyber, does not prove to be excessive or disproportionately ruinous, given the force of the initial attack. The rough idea, perhaps, is to the balanced scales of Lady Justice herself: the reply may not be grossly over-weighted beyond the quantum of the initial attack (and, maybe further, what is needed for reliable security from an imminent repeat of such an attack).

Rule 73 offers further criteria governing the circumstances within which a state may utilize force in the interests of self-defense.[134] Here the Experts deviate slightly from the explicit wording of Article 51 to make provisions for armed cyberattacks which have not yet occurred

[130] *Tallinn 2.*0, 347.
[131] *Tallinn 2.0*, Rule 72: "A use of force involving cyber operations undertaken by a State in the exercise of its right of self-defence must be necessary and proportionate."
[132] *Tallinn 2.0*, 349.
[133] *Tallinn 2.0*, 349.
[134] *Tallinn 2.0*, Rule 73: "The right to use force in self-defence arises if a cyber armed attack occurs or is imminent. It is further subject to a requirement of immediacy."

but are considered imminent.[135] The Experts argue that the imminence requirement is met only in the event of another state making preparations or expressing clear intent to launch an attack; without reasonable belief that an attack is forthcoming, a potential victim state remains limited to acts short of force to defend or dissuade against the (mere) machinations of another state.[136] The corresponding principle of immediacy seeks to ensure that forceful reactive action is taken with the intent of self-defense, rather than merely retaliation.[137] This principle demands that any use of force taken in response must do so only in a window immediately following the initial attack, within which the threat of continued cyber operations remains. Should a cyberattack cease, the victimized state is unjustified in employing force in its response, provided they are aware that the belligerent cyberoperation is unlikely to resume.[138] Much like the use of force and armed attack distinction, this is likewise reliant on a state's reasonable judgement rather than an explicit set of guiding criteria.

Finally, Rules 74 & 75 briefly expand on provisions within Article 51. Rule 74 extends the right to collective self-defense even in the event of a cyberattack, contingent on the request of the victim state.[139] Estonia, obviously, would strongly endorse such—and wished that NATO had agreed, in a more timely way, back in 2007. Meanwhile, Rule 75 directly extends the requirements originally outlined within Article 51, mandating that all measures taken in the name of self-defense be reported immediately to the UN Security Council.[140] Both Rule 75 and Article 51 serve to indirectly fulfill the *jus ad bellum* requirement of proper declaration. In demanding that retaliatory measures be reported to the Security Council, both Rule 75 and Article 51 work towards ensuring that responses in self-defense are clearly attributable and that the victim's grounds for forceful response are clearly conveyed. Rather than simply being a courtesy, this reporting requirement serves two crucial functions for conflict resolution. First, it ensures that the UN Security Council maintains arbitrational authority over the conflict, including the

[135] *Tallinn 2.0*, 350.
[136] *Tallinn 2.0*, 353.
[137] *Tallinn 2.0*, 353.
[138] *Tallinn 2.0*, 353.
[139] *Tallinn 2*.0, Rule 74: "The right of self-defence may be exercised collectively. Collective self-defence against a cyber operation amounting to an armed attack may only be exercised at the request of the victim State and within the scope of the request."
[140] *Tallinn 2.0*, Rule 75: "Measures involving cyber operations undertaken by States in the exercise of the right of self-defence pursuant to Article 51 of the United Nations Charter shall be immediately reported to the United Nations Security Council."

authority to issue a stand down order to the responding state, provided the Security Council has implemented its own conflict resolution measures.[141] Secondly, and perhaps more importantly in the case of cyberwar, the mandatory reporting of retaliatory measures tries to ensure that they are attributable. One of the fears surrounding retaliation strikes is the potential for these retaliations to be misattributed to another uninvolved state. This is a particular concern in the case of cyberattacks which are conducive to misdirection efforts—complex routing patterns, cloaked "Trojan horse" malware viruses, etc.—and thus benefit from plausible deniability. Rule 75 ensures that measures launched in self-defense are done so by the proper authority, against an appropriate target, and without misdirection which may mask the retaliation as an opportune attack by another state, further intensifying the conflict.

Accordingly, the *Tallinn 2.0* rules pertaining to cyber sovereignty, digital uses of force, and armed cyberattacks offer a strong basis for a robust evaluative legal framework covering each of the four basic *jus ad bellum* principles, as straightforwardly understood by, and specified within the LOAC. Rules 1-4 offer a defense of sovereignty within cyberspace, demarcating three layers of cyberspace to which states may reasonably lay claim. Working together with the provisions laid out in Article 2 of the UN Charter, these rules offer a pathway towards meeting the just cause condition of *jus ad bellum* via cyber operations; should a cyber operation prove a grievous violation of a state's sovereignty, the UN may feasibly be justified in deploying force once diplomatic options have been exhausted to no avail. A secondary means of meeting the just cause principle is covered by Rules 69-75, with the Experts' deliberations suggesting that cyber operations are readily capable of rising to the level of an armed attack. In the event of a use of force crossing the threshold of an armed attack, Article 51 may be invoked, and the victim state would be justified in responding with force of their own, irrespective of whether the origin of the attack was physical or digital. As a result, *Tallinn 2.0* offers two specific circumstances within which the just cause principle may be met with regards to cyber operations.

With the "prime mover" of the just cause principle covered, the remaining three LOAC *jus ad bellum* principles (of last resort, proportionality, and proper declaration) are likewise addressed within Rules 72 and 75. The first of these rules asserts that retaliatory uses of force only prove justified once other less forceful means have been exhausted or deemed insufficient

[141] *Tallinn 2.0*, 356.

for defending against an attack. This principle works towards mitigating unnecessary uses of force potentially resulting in an escalation of hostilities. Rule 72 further demands that any retaliatory force employed must be proportionate to the harm incurred; this constrains the level of force available to a victim state, preventing it from causing catastrophic levels of harm in response to an attack of significantly lesser scale. Finally, Rule 75 introduces the principle of proper declaration through the final provision of Article 51, asserting that any action undertaken in the interests of self-defense must be reported in a timely manner to the UN Security Council so that there can be accountability as well as consideration of any required collective action. Taken holistically, these two sets of rules within *Tallinn 2.0* offer a strong governing framework for both the evaluation of incoming cyberattacks, as well as for gauging the permissibility of potential responses in the interest of self-defense. As a result, *Tallinn 2.0* undoubtedly represents a strong and substantial effort at redressing the present sense of lawlessness which permeates the cyber domain through the extension of presently binding international law.

3.4 Pitfalls for the Tight-to-the-Law Approach

Unfortunately, hopes that the *Tallinn Manual* would motivate, if not shape, international efforts towards cyberwar governance were quickly dashed, and seem to remain that way for the foreseeable future. Despite its impressive work towards the interpretation of existing LOAC and how they may apply to cyberspace, the document has hitherto enjoyed little endorsement within the broader international community. The general provisions contained within the *Tallinn Manual*'s rules have likewise had seemingly little influence on how states plan and deploy cyber operations, resulting in the same general fraught equilibrium, described in the previous chapter, which had existed prior to the drafting of the document. The lack of acknowledgement received by the *Tallinn Manual* has led to some UN and adjacent NGO officials to go so far as to declare the overall effort indeed to be "a spectacular failure".[142] Evidently, the rules offered by the *Tallinn Manual* have failed to persuade the global community, despite its appeals to presently

[142] George R. Lucas Jr., "Emerging Norms for Cyber Warfare," In *Binary Bullets: The Ethics of Cyberwarfare*, eds. Fritz Allhoff, Adam Henschke, & Bradley Jay Strawser (New York, NY: Oxford University Press, 2016), 17.

binding LOAC which have been agreed to, ratified, and largely adhered to by the member states of the UN.

In the eyes of its detractors, the *Tallinn Manual* faces challenges along two specific axes. First, the tight-to-the-law approach undertaken by the document leads some to accuse the *Tallinn Manual* of "jurisdictional equivocation".[143] While treating the cyber domain as fundamentally akin to the traditional domains of conflict renders it possible to position cyber operations under the purview of existing international law, it also inadvertently leaves the *Tallinn Manual*'s findings vulnerable to criticism by opponents arguing that the cyber domain is instead fundamentally different to the traditional domains. Although the Experts had rejected a conception of cyberspace as a form of global common, stakeholder nations such as China nonetheless insist—for example—that the cyber domain represents a public space, and thus carving it up into parcels of national sovereignty misconceives the nature of the space.[144] While this disagreement may be in part credited to the adversarial relations between the US and China, it is worth noting that the *Tallinn Manual* team, being a NATO-led initiative, excluded representation from major cyberpower states such as China and Russia.[145] Accordingly, it is not unexpected that the legal interpretations of Western Experts would generate scrutiny and controversy amongst UN members whose input had been shunned in the drafting process. As a result, the tight-to-the-law approach, coupled with the exclusion of relevant stakeholder input, apparently ensures that the *Tallinn Manual* will enjoy limited support amongst non-NATO states.

The shortcomings of the *Tallinn Manual* approach are perhaps best evidenced by comparisons to other efforts at "building consensus for future governance", such as the 2008 Montreux Document.[146] The Montreux Document was an international effort towards regulating the usage of private military companies (PMCs) resulting from deliberations between relevant stakeholders including not only nations such as China, but also PMC representatives.[147] Much like the *Tallinn Manual*, the Montreux Document does not present itself as a legally binding article; instead, it seeks to reiterate pertinent legal obligations drawn from international law *as*

[143] Lucas, "Emerging Norms for Cyberwarfare", 19.
[144] Lucas, "Emerging Norms for Cyberwarfare", 20.
[145] Lucas, "Emerging Norms for Cyberwarfare", 17.
[146] Lucas, "Emerging Norms for Cyberwarfare", 17.
[147] Lucas, "Emerging Norms for Cyberwarfare", 17.

well as to develop norms of best practice for the usage of PMCs.[148] Lucas credits this resulting "soft law" approach as motivating the Montreux project towards greater international endorsement than the stricter legal approach of the *Tallinn Manual* has been able to generate.[149] By contrast, the *Tallinn Manual*'s tight legal approach, coupled with the exclusion of the input of other relevant cyberpowers in the deliberation process, hampers its ability to garner the degree of acknowledgement necessary for it to meaningfully influence global cyber governance.

The second difficulty facing the *Tallinn Manual* stems from the inherent shortcomings of the analogy approach when evaluating the permissibility of cyberattacks resulting in non-physical or unique harms. Although the *Tallinn Manual* proves largely successful in logically extending LOAC into cyberspace, unique elements of cyber operations ensure that this transition is neither seamless nor complete. Much of the *jus ad bellum* work done within the document is built upon the presumed similarities between cyber operations and their kinetic cousins. This is made most evident by the Experts' conclusions regarding the scale and effects considerations of cyberattacks; these evaluations are reliant on the harms of a given cyberattack being analogous to the harms of conventional operations which the international community readily recognizes as uses of force and violations of sovereignty. For example, an act of cyber sabotage destroying physical infrastructure is uncontroversially regarded as a use of force on the grounds that its scale and effects resemble what might be achieved by conventional means. In the same vein, cyberattacks can rise to the level of armed attacks despite their digital origins, provided their scale and effects are sufficiently severe to be comparable to those of conventional armed strikes.

Although this approach through analogy offers a great first step towards the evaluation of cyber operations, it nonetheless cannot adequately account for the full spectrum of cyberattacks. While cyberattacks resulting in lasting physical damages occasionally occur, they are comparatively rare. By comparison, many more cyberattacks fail to cross the digital-physical divide—at least in a direct, violent manner. The NotPetya attacks which paralyzed Ukrainian infrastructure were primarily digital; despite the worm proving capable of turning off Western Ukraine's power, it did so without causing tangible damage to grid infrastructure itself. Likewise, the DDoS attacks on Estonia were wholly digital in nature, working to disrupt

[148] International Committee of the Red Cross, *The Montreux Document*, (Geneva: International Committee of the Red Cross, 2009), 9.
[149] Lucas, *"*Emerging Norms for Cyberwarfare*"*, 18.

Estonian cyber infrastructure without causing permanent physical damages; while it is arguable that harms occurred, in the form of Estonians being unable to access a range of digital services, these harms are of a materially different nature to those incurred as a consequence of conventional action. More recently, a cyberattack on the German district of Anhalt-Bitterfeld managed to functionally paralyze the region's digital services, causing the district to be the first in Germany to formally declare disaster due to a digital attack, despite the absence of physical damage.[150] In each case, a cyber operation was able to achieve its objective without causing permanent real space damages. Should a conventional operation have achieved the same result, perhaps by physically destroying hardware integral to a state's information networks, it would likely have been labelled a use of force, or perhaps an armed attack, due to the readily quantifiable element of physical damage as well as the corresponding financial and material costs of repairing such damages; however, in the absence of comparable physical harms, such declarations become far more controversial. When it is just a "pause in functionality" which can be (or even is) deliberately reversed, then it does seem to be a different, non-analogous, kind of action.

Non-physical harms wrought by cyber operations generally pose problems for the *Tallinn Manual*'s evaluation efforts both in terms of potential violations of sovereignty, as well as the use of force and armed attack distinctions. The Experts unanimously agreed that cyberattacks resulting in physical damages, including those that destroy physical elements of a target's cyber infrastructure, would constitute a violation of the victim's sovereignty "because such consequences are akin to physical damage or injury".[151] The Experts would further share the sentiment that cyber operations resulting in a loss of functionality of another state's cyber infrastructure could *sometimes* be regarded as violations of sovereignty; the distinction as to *where* this threshold lies would remain unresolved however, as the Experts were unable to unearth any relevant legal precedents to properly ground this conclusion as a rule.[152] Evidently, further problems with the analogy approach would emerge as the nature of the harm incurred grows less recognizable in conventional terms.

---

[150] "Rural German District Declares Disaster After Cyberattack", *Deutsche Welle*, July 10, 2021, Accessed August 8, 2021, https://www.dw.com/en/rural-german-district-declares-disaster-after-cyberattack/a-58227484.
[151] *Tallinn 2.0*, 20-21.
[152] *Tallinn 2.0*, 20-21.

Yet more uncertainty revolves around the evaluation of cyberattacks which fail at all to result in physical harms or losses of functionality. With regards to these cases, the Experts were divided as to whether such operations could qualify as violations of sovereignty. One contingent of Experts, believing that non-physically harmful cyberattacks could constitute such violations, argued such operations could serve to actively undermine the principle of sovereignty; attacks such as large-scale disruptive DDoS operations and the remote alteration of data were identified by these Experts as clear violations of a state's right to full control over cyber activities within its territory.[153] This interpretation failed to garner consensus amongst the rest of the Experts, however. While the Experts were unanimous in their agreement that cyber operations undermining a state's ability to perform governmental functions, such as elections or the distribution of welfare, were unlawful violations of a state's sovereignty, it was more difficult to distinguish other non-physically harmful cyber operations as likewise constituting such violations.

The *Tallinn Manual*'s analogy approach also proves particularly problematic when evaluating whether cyber operations rise to the level of uses of force or armed attacks. The destruction of a military installation is considered a use of force irrespective of whether the attack was a conventional airstrike or an explosion caused remotely by digital means; despite having digital origins, such a cyber operation would result in the kinds of familiar physical harms which have regularly been recognized by the international community as being sufficient for declaring that a use of force has been committed. In these cases, the analogy approach is straightforwardly sufficient. However, it proves difficult, if not impossible, to find neat analogues for many offensive cyber operations. Disruptive cyberattacks are theoretically capable of functionally shutting down an entire state until the state's defensive entities can work out a solution or the belligerent party opts to cease its attack. While it is imaginable that a conventional attack may achieve the same disruptive effect, it is unlikely to do so without causing significant physical damages to the relevant infrastructure, which would likely serve to escalate the situation to full blown war. Accordingly, the evaluation of severe disruptive digital attacks cannot be appropriately measured with comparisons to kinetic operations as significant discrepancies exist between cases; while both kinds of operations achieve a similar goal, they do

[153] *Tallinn 2.0*, 21.

so in drastically disparate manners resulting in different scales and effects, rendering the two cases at least somewhat disanalogous. Consider, for example, a cyber case where electricity is shut down for several hours but can be, and is, booted up again, versus a kinetic case where the electricity is shut down because the infrastructure gets destroyed in an airstrike. In the latter, repair efforts are economically costly and time consuming, with power potentially being down for weeks or months. By contrast, the damages in the cyber case may very well prove entirely reversible with merely a few keystrokes.

The analogy approach faces further difficulties when considering cyberoperations resulting in unique harms beyond the scope of what is attainable via conventional means. In his analysis of cyber conflict, Lucas asserts that cyber operations can be broken into two specific forms depending on the nature of their effects: *physical effects-based* and *political effects-based*.[154] Physical effects-based cyber operations encompass those cyberattacks which result in the sorts of familiar physical harms which are readily perceived as being analogous to conventional operations, such as the physically destructive Stuxnet worm which shut down an Iranian nuclear reactor and could have, if undetected, even resulted in a meltdown scenario.[155] By contrast, Lucas conceives of political effects-based attacks as consisting of cyber operations designed explicitly to "*impose the cyber aggressor's political will upon its adversaries through nonpolitical means*".[156] This would include operations of the sort launched by Russia against Ukraine in the wake of the Euromaidan unrest, within which Russian-backed operators sought to sow discord and shape domestic politics in support of a manufactured pro-Russian narrative. Accordingly, these kinds of operations seek to project foreign policy goals without relying on political discourse, economic warfare, or physically harmful attacks. Far from being rare occurrences in a digital battleground dominated by physical effects-based cyber operations, Lucas argues that these kinds of political effects-based attacks are more indicative of the future of interstate conflict than their physically destructive counterparts.[157] As states largely avoid the usage of destructive cyber weapons with the potential for incurring severe retaliation, they turn

[154] Lucas, *Ethics and Cyber Warfare*, 57.
[155] Lucas, *Ethics and Cyber Warfare*, 57.
[156] Lucas, *Ethics and Cyber Warfare*, 9.
[157] Lucas, *Ethics and Cyber Warfare*, 9

towards softer cyber alternatives capable of achieving their foreign policy objectives with less corresponding undue risk (albeit over longer time-frames).

The prevalence of political effects-based cyberattacks serves to undermine the Experts' analogy approach by distancing the predominant kinds of cyberattacks from the sort with readily comparable kinetic analogues, rendering it difficult to ascertain whether one such operation has constituted a use of force. In keeping with the precedent set by the UN's Declaration on Friendly Relations in 1970, the Experts assert, in *Tallinn 2.0*, that (mere) political and economic pressure falls short of being regarded as uses of force; this precedent further motivated the Experts to regard non-destructive psychological operations, seeking to undermine confidence in a state, as similarly falling short of uses of force.[158] While these declarations may have sufficed for the kinds of operations that were possible in the 1970s, the advent of cyberweapons and the general trend of ever greater state reliance on information technology necessitates a rethink of the potential of political effects-based cyber operations. Presently, cyber operators enjoy a much more active role in undermining a populace's confidence in their government than was possible for intelligence operators in the 1970s. Cyber actors have proven capable of remotely controlling the information sphere through the targeted suppression of information from one side of a conflict while tactically promoting the views of the other. Widespread but temporary disruptions to integral civilian infrastructure, such as banking and healthcare, have likewise proven to have devastating effects on a state's domestic politics, all without necessarily causing permanent physical harms. As seen in the Ukraine conflict, these tactics have undeniable efficacy well beyond what is achievable through conventional psychological operations reliant on traditional state propaganda outlets, political posturing, and support for local dissidents. Accordingly, the invasiveness and effectiveness of modern cyber operations warrants greater discourse regarding their status as potential uses of force, irrespective of whether they result in physical damages.

Finally, the analogy approach to the evaluation of cyber operations faces difficulties in ascertaining when a cyber use of force crosses the threshold of becoming an armed attack sufficient for the purposes of Article 51. Here the Experts in *Tallinn 2.0* once again proved unable to achieve consensus regarding the necessary criteria to be met. Some argued that the presence of physical harm or destruction remains a prerequisite for an armed attack, while others

[158] *Tallinn 2.0*, 331.

pointed towards the scale of the effects rather than the nature of an attack's consequences.[159] The Experts were similarly divided in the hypothetical case of a cyber operation against an international stock exchange. The Experts against labelling this kind of operation as an armed attack were staunch in their beliefs that "mere financial loss" was insufficient for qualifying such an attack as being armed; conversely, other Experts argued that the losses incurred from such an attack could feasibly prove to be devastating enough to distinguish the associated operation as being an armed attack.[160] In the absence of legal precedents offering an explicit threshold for cyber uses of force and cyber armed attacks, even the case of Stuxnet proved controversial as the Experts, while in consensus that Stuxnet represented a use of force, remained split as to whether the worm represented an armed attack.[161] Evidently, the unique nature by which Stuxnet had achieved its objectives proves sufficient for distinguishing it from conventional tactics employed towards the same ends. Once again, relying on the existence of analogues to distinguish armed attacks from mere uses of force proves insufficient for the evaluation of cases within which the methodology of a cyberattack drastically differs from traditional kinetic alternatives.

Evidently, while cyber operations may not offer something entirely new in the field of conflict, they nonetheless offer something *different*, whether due to their effects, or to the means with which they achieve their objectives, or to the ease with which some of their effects can be undone (with some infrastructure) or not (with some of the political psychology manipulation). While the *Tallinn Manual* provides a compelling evaluative framework extending existing LOAC into cyberspace, its ability to account for cyberattacks resulting in nonphysical harms is hindered by its reliance on the existence of analogues. Cyberattacks resulting in large-scale conventional harms such as physical injury, death, or destruction, have been comparatively few and far in between —though they may increase— with the majority of cyberattacks thus far failing to transcend the physical-digital divide. Insofar as these latter kinds of attacks are often launched with objectives disparate from conventional operations (such as an emphasis on temporary disruption rather than physical destruction) while simultaneously employing unique methodologies for achieving these objectives, it appears increasingly unlikely that it will be possible to find apt analogues or legal precedents for the entire spectrum of cyber operations.

[159] *Tallinn 2.0*, 342.
[160] *Tallinn 2.0*, 342.
[161] *Tallinn 2.0*, 342.

Moreover, the advent of cyber tactics has necessitated a rethink of how political influence operations are conducted; information warfare operations have grown to be more direct, invasive, and potent with the widespread adoption of network technologies. Despite non-violent political pressure having been regarded as failing to constitute a use of force throughout the 20th century, modern cyber tactics serve to render contemporary political pressure campaigns disanalogous to those conducted before the age of unprecedented connectivity, suggesting that the earlier dismissal of such operations as potential uses of force warrants reconsideration. Accordingly, while the *Tallinn Manual* represents an excellent effort towards governance within the cyber domain, it struggles to adequately account for disanalogous types of cyber operations, leaving a significant gray area within its evaluative framework, indeed as highlighted by the noted disagreement amongst the Experts themselves.

3.5 Hotfixing Cyberwar Governance

While the Montreux Document bears some similarities to the *Tallinn Manual*, it seems—remarkably—to avoid the pitfalls which plague the latter document in part due to its broader scope. Rather than focusing purely on potentially pertinent LOAC, the Montreux Document adopts a more unfettered, straightforwardly normative approach, seeking to engage with relevant stakeholders to identify best practices regarding the employment of PMCs, before disseminating the conclusions to the international community for further deliberation.[162] To this end, the Montreux Document is broken into two main parts. The first of these contains the relevant legal obligations pertaining to PMCs. Echoing the work of the *Tallinn Manual*, the first section of the Montreux Document does not introduce new laws, but rather seeks to reiterate LOAC and customary international agreements deemed relevant for the governance of PMCs.[163] However, the second section of the document serves to outline various norms of best practice which it urges states to consider in their decision making regarding the employment of PMCs.[164] While these norms are (understandably) not legally binding, they serve two important roles in the decision-making process. Firstly, the norms outlined in the second section encourage states to act

[162] Lucas, "Emerging Norms for Cyberwarfare", 18.
[163] *The Montreux Document*, 13.
[164] *The Montreux Document*, 16.

in a manner which preserves human rights. Secondly, and perhaps more importantly, the norms provided offer states a flexible set of criteria and guidelines which can be applied to a wide spectrum of situations, rather than merely those which fall neatly into international law or those that are analogous to past cases.

It is worth stressing further that sometimes, with very controversial issues in international relations, it can be the better part of wisdom to go for the informal, so-called "soft law" (or merely advised practices) approach. Going straight-away for the binding "hard law" of ratified treaties can, so to speak, scare states away from the whole enterprise. The stakes become too high, too fast. States may prove reluctant to agree and adhere to binding legal standards with readily enforceable punitive measures. This results in no norms at all, which is arguably what has happened thus far with cyber-conflict. I have already argued, at the start of last chapter, why having some norms here is much better than none at all. Further, the informal normative approach can be more welcoming and inclusive, thus perhaps drawing in the participation of states whose involvement—like it or not—is crucial to the viability and endurance of any normative regimes.

The comparative success of the Montreux Document in fostering dialogue and agreement amongst stakeholders suggests that a more normative approach is necessary for governing conflict in the cyber domain. While the narrow legal focus of the *Tallinn Manual* renders it difficult to evaluate the full spectrum of cyber operations, due to a lack of conventional analogues for cyberattacks resulting in unique harms, the flexibility of normative criteria ensures that they remain applicable in all cases. Consequently, I argue that the legal *jus ad bellum* as prescribed by the *Tallinn Manual* ought to be supplemented with considerations present within the *jus ad bellum* of the Just War Theory (JWT) tradition. While the LOAC discussed within the *Tallinn Manual* provide a strong framework for cyber operations with physical analogues, the normative considerations of JWT offer further guidance not only for analogous cases, but also for those cyber operations which either result in unique harms or those whose methodology is sufficiently different to render them unfamiliar territory. Taken in tandem, what results is a robust evaluative framework for cyber operations which is not only comprehensive, but readily adaptable for the analysis of a wider range of cyber operations. Furthermore, by virtue of distancing itself from controversial legal declarations, a JWT-based framework for cyberwarfare would instead build on the shared interests of relevant stakeholders; this allows for a broader

evaluative framework within which we may begin identifying emerging best practices and the values which ought to govern cyberoperations.

While the LOAC specifies four criteria for fulfilling *jus ad bellum* (namely, just cause, proportionality, last resort, and public declaration), just war theory frameworks typically identify two further criteria in the form of right intention and probability of success.[165] The following chapters will be comprised of a close analysis of each of these six JWT principles to show how, from this more informal and expansive moral framework, each principle translates into the cyber domain. This will be done to build a comprehensive cyber *jus ad bellum* for the evaluation of cyber attacks and to define the circumstances under which each of the principles may be satisfied. For organizational purposes, these six principles will be divided into two larger, overarching categories: consequentialist principles and anti-consequentialist principles.[166] The subsequent two chapters will cover the anti-consequentialist "first principles" of just cause, right intention, and public declaration. The remaining three consequentialist principles of proportionality, last resort, and probability of success will then be covered in a single chapter. The resulting cyber *jus ad bellum*, while not itself legally binding, will offer a strong conceptual and ethical foundation for the evaluation of contemporary cyberattacks and a rational basis for compelling discourse regarding best practices amongst cyberpowers.

[165] Orend, *War and Political Theory*, 89.
[166] Orend, *War and Political Theory,* 92-93.

Chapter 4

Just Cause

4.1 Rethinking Just Cause for the Cyber Context

Despite the *jus ad bellum* of JWT being holistic in nature, being met if and only if each of its six criteria are fulfilled, the criterion of just cause enjoys particular importance amongst its peers. This significance stems from the just cause criterion serving as the first determining factor for gauging whether a state's military action, be it a reaction on grounds of self-defense or a proactive humanitarian intervention abroad, is morally justified.[167] Within the context of self-defense, the immediate aftermath following an aggressive action waged against a state by another forces the targeted state to discern whether the incurred act of aggression satisfies this *jus ad bellum* criterion. Should the condition of just cause be deemed met by the aggressive action, the targeted state may prove justified in responding in kind with their own deployment of force. However, if the incursion fails to fulfill the just cause criterion, the responses available to the targeted state are limited to short-of-force operations. Further deliberations regarding responses to aggression, such as calculations of proportionality for proposed retaliatory measures, are necessarily predicated by this initial assessment of the incurred act of aggression; in the absence of a satisfied just cause criterion, it is ultimately irrelevant whether the targeted state's response is proportional or conducted as a matter of last resort as the state remains fundamentally *unjustified* in responding with force of their own. The usage of force in the absence of just cause would render the responding state open to sanction or other penalty by international organizations, and perhaps even mark itself as a target for subsequent retaliation at the hands of the initial aggressor state, now with just cause of its own. Accordingly, the criterion of just cause serves as the prime mover of the overall *jus ad bellum* project, being the first criterion which must necessarily be fulfilled prior to opening considerations of the remaining five.

[167] Brian Orend, *War and Political Theory*. (Cambridge, UK; Medford, MA: Polity, 2019), 86.

Insofar as the just cause criterion motivates the *jus ad bellum* project and distinguishes permissible from impermissible kinds of reactions to force, the task of identifying the relevant threshold at which this criterion is met takes on utmost importance. A threshold set too low significantly broadens the spectrum of attacks which would qualify as reasonable grounds for going to war. This would potentially loosen restrictions on states' usage of force as more types of belligerent actions satisfy the just cause condition. This may result in an overly cavalier approach to war within which even minute acts of aggression may justifiably escalate into full-blown conflict. In contrast, setting the threshold too high presents the opposite problem. Should the just cause criterion be considered met only in the face of the most severe acts of aggression, we run the risk of targeted states finding themselves incapable of forceful deterrent unless they are currently being targeted with a concerted military offensive or undergoing an armed invasion, in which case it may be too late for any such responses to prove effective. A higher threshold may actually serve to embolden aggressors, allowing them to operate with a greater degree of impunity owing to the knowledge that their targets would prove unjustified in responding to aggression with force in all but the most severe of cases. As a result, the first step into a cyber *jus ad bellum* requires establishing a balanced and proper threshold which proves neither too lenient so as to prevent the just cause criterion from becoming trivial, nor too stringent so that only the worst acts of aggression, such as the so-called cyber-Pearl Harbour predicted by doomsayers,[168] motivate the principle of just cause.

As discussed within the preceding chapter, the United Nations Charter and the *Tallinn Manual* offer two distinct means by which the just cause criterion may be fulfilled. In the first instance, both acknowledge that grievous violations of state sovereignty can be conducted through purely digital means, and that such actions may be deemed sufficient by the UN for the deployment of force once non-forceful alternatives have been exhausted. Notably, this form of justification has a multilateral element to its deliberative process; it is ultimately the UN which determines whether the just cause criterion has been met, and it is the UN which coordinates any corresponding forceful or short-of-force responses to aggression. In the second instance, the just cause criterion is fulfilled by the immediate need for self-defense which is triggered by being

[168] Elisabeth Bumiller and Thom Shanker, "Panetta Warns of Dire Threat of Cyberattack on U.S.," *The New York Times*, October 11, 2012. Accessed January 15, 2022. https://www.nytimes.com/2012/10/12/world/panetta-warns-of-dire-threat-of-cyberattack.html.

targeted by an act of aggression rising to the level of an armed attack. In these latter cases, the just cause criterion authorizes a targeted state to respond unilaterally, potentially with force should other defensive measures be exhausted or otherwise deemed inadequate for deterring further aggression. Under both motivating circumstances, the just cause threshold is set fairly high in the case of cyberwar. Only the strongest acts of cyberaggression are likely to garner the multilateral consensus necessary for asserting that a state is justified in going to war. Multilateral agreement is further rendered less likely due to the general controversy surrounding concepts such as ownership and territory within cyberspace. The armed attack distinction in the cyber context proves similarly controversial, increasing the difficulty of distinguishing the cases within which states are justified in unilaterally responding to severe cyberattacks with force of their own.

Although the existing threshold for the just cause criterion outlined within the *Tallinn Manual* may serve to prevent comparatively mild cyberattacks from serving as impetus for an escalation of hostilities, they nonetheless prove too restrictive when considering the full spectrum of cyber operations. Specifically, deliberations regarding when an armed cyberattack has occurred place too strong an emphasis on the presence of physical damages or bodily harm directly attributable to the attack; barring attacks resulting in harms analogous to conventional action, it remains highly unlikely that the unilateral self-defense justification will obtain in the case of purely cyber operations. This becomes problematic insofar as the vast majority of cyberattacks fail to result in these kinds of harms. Rather, most cyberattacks proceed along two distinct trajectories which render them disanalogous to conventional alternatives. Firstly, many cyberattacks are launched with objectives beyond the capabilities of conventional action, such as information theft and targeted disinformation campaigns. Secondly, those cyberattacks launched towards familiar objectives, such as the destruction of stockpiled munitions, arrive at their objectives in a materially different fashion, precluding them from being considered equivalent to kinetic alternatives. In the case of Stuxnet, the remote crippling of a controversial nuclear enrichment facility was treated by all parties as fundamentally different from an air strike achieving the same result. What results is a lopsided digital battlefield within which aggressors not only benefit from the latent advantages cyber offenses enjoy over cyber defenses, but also the general difficulty defenders face in motivating strong responses to offensive cyber operations.

The unique nature of disanalogous cyberattacks warrants rethinking the current thresholds for just cause within the cyber context. As evidenced within the preceding chapters, the potential damages wrought by offensive cyberattacks are not limited to the kinds of consequences which may be incurred as a result of kinetic action. While the *Tallinn Manual* offers a compelling means of evaluating cyberattacks resulting in the "scale and effects" traditionally associated with conventional action, I argue that we require a more flexible evaluative approach towards non-physically harmful cyberattacks to determine whether such operations may nonetheless qualify as acts of aggression sufficient for fulfilling the just cause criterion. To this end, the following section seeks to examine the genesis of the just cause criterion from the initial concept of state sovereignty and the corresponding rights and duties it bestows upon the state. This analysis will then be developed to extend the original intent behind the inherent right to self-defense into cyberspace, showing that non-physically destructive cyberattacks underming state sovereignty ought to be seen as having the potential to motivate justified responses in the interest of self-defense, even in the absence of familiar physical harms.

4.2 The Heart of the State and the Right to Self-Defense

The concept of aggression lies at the heart of every dialogue regarding state self-defense. The history of the JWT tradition suggests that defensive wars fought in reaction to incurred aggression are more straightforwardly morally justifiable than offensive ones. If a state finds itself besieged by a belligerent neighbour, it is uncontroversial to assert that the beleaguered state is well within its rights to respond to the aggressor with force. In most such cases, the targeted state's response is not only accepted, but supported, either tacitly or explicitly, by other members of the international community who recognize that the aggressor has violated some semblance of harmony and order amongst states. Evidently, there is something inherent to the act of interstate aggression which morally justifies forceful resistance.

In pursuit of explaining *why* this is, Walzer posits that aggression is "the only crime that states can commit against other states", existing as a capital offense without the degrees of severity that we typically ascribe to interpersonal crimes, such as the legal distinction between

murder and manslaughter.[169] Our knowledge of this crime, according to Walzer, stems from our familiarity with peace in its absence; notably, this is not merely peace demarcated by a lack of open warfare between states, but rather a stronger peace Walzer terms "peace-with-rights" within which aggression itself is absent and citizens' rights are protected.[170] Acts of aggression serve to violate this peace-with-rights, posing an existential threat to the rights of citizens and forcing them to choose between their rights and their personal safety.[171] For Walzer, the coercion of this choice *always* justifies a forceful response on behalf of those forced into this decision by an aggressive action. Accordingly, aggression morally authorizes resistance by virtue of the threat it poses to the rights of citizens and, by extension, states.

This leaves open the question of which inherent rights are enjoyed by political communities, moves against which would constitute acts of aggression sufficient for fulfilling the principle of just cause. For Walzer, there are two specific rights integral to the state: territorial integrity and political sovereignty.[172] Crimes of aggression consist of foreign states violating one or both of these rights, either through the deployment of military force across borders, such as any of countless territorial invasions recorded throughout human history, or through subtler direct and malign interference into the political deliberations of foreign states, perhaps best evidenced by the Central Intelligence Agency's efforts to stem the growth of socialism throughout South and Central America throughout the Cold War, with tactics ranging from solidifying the rule of neofascistic juntas, such as that of Pinochet in Chile in the 1970's, with financial or intelligence support,[173] to outright arming and training opposition paramilitaries such as the Nicaraguan Contras in the 1980s.[174] In both cases—and any similar ones— the state whose rights are being violated by external interference may be justified in motivating a forceful multilateral or unilateral response per provisions in Articles 42 & 51 of the UN Charter.

---

[169] Michael Walzer, *Just and Unjust Wars: A Moral Argument with Historical Illustrations,* 4th ed. (New York: Basic Books, 2006). 51-52.
[170] Walzer, *Just and Unjust Wars*, 51.
[171] Walzer, *Just and Unjust Wars*, 51.
[172] Walzer, *Just and Unjust Wars,* 53.
[173] Kevin A. O'Brien, "Interfering with Civil Society: CIA and KGB Covert Political Action during the Cold War," *International Journal of Intelligence and CounterIntelligence* 8, no. 4 (December 1995): 431–56, https://doi.org/10.1080/08850609508435297. 438.
[174] *Case Concerning Military and Paramilitary Activities In and Against Nicaragua (Nicaragua v. United States of America); Merits*, International Court of Justice (ICJ), 27 June 1986.

While these rights are often considered the exclusive domain of world states, Walzer asserts that these two rights which underpin international relations "derive ultimately from the rights of individuals, and from them they take their force".[175] Accordingly, interstate aggression cannot be regarded as a crime conducted in a vacuum between two abstract entities. Rather, for Walzer, the harms perpetrated by interstate aggression are most palpably felt by the citizens comprising the state. At the outbreak of hostilities, it is the citizens' lives which come under fire; by extent, it is the values of the citizens, such as their political associations, which are assailed by the military efforts of a foreign belligerent.[176] The grave nature of aggression stems from its deliberate undermining effect on the political self-determination of a collective of individuals. The right of state self-determination is not fundamentally different from the rights of the citizens; rather, Walzer conceptualizes states' rights as a collective form of citizens' individual rights, with the state's political self-determination itself reflecting the self-determination—the freedom and autonomy—of its citizens.[177]

Although Walzer asserts that state rights are contingent on the consent of their citizens, he does not conceive of this consent as an explicit series of rights transfers in exchange for the protection that membership of a state affords. Instead, Walzer believes that generations of shared life experiences and close cooperation amongst individuals eventually leads to the manifestation of a "common life" which resides at the heart of all states.[178] For Walzer, the notion of a community goes much deeper than merely being a collective of individuals bound by the confines of territorial borders. Rather, Walzer asserts that it is within political communities that "[l]anguage, history, and culture come together (come more closely together here than anywhere else) to produce a collective consciousness".[179] A history of shared experiences leads to the development of a community character aligned with the emergent sensibilities and values guiding both the collective's internal deliberations and its interactions with external political communities.[180]

---

[175] Walzer, *Just and Unjust Wars*, 53.
[176] Walzer, *Just and Unjust Wars*, 53-54.
[177] Walzer, *Just and Unjust Wars,* 54.
[178] Walzer, *Just and Unjust Wars*, 54.
[179] Michael Walzer, *Spheres of Justice: A Defense of Pluralism and Equality*. (New York: Basic Books, 2010). 28.
[180] Walzer, *Spheres of Justice*, 28.

This is not to say that communities are homogenous entities with immutable values and sensibilities. Walzer notes that it is regularly the case that smaller collectives within the greater community of the state, such as distinct regional or cultural groups, have sensibilities which may diverge from those of the state as a whole.[181] These differences, Walzer reasons, are resolved politically through deliberations between these smaller collectives based "upon understandings shared among the citizens about the value of cultural diversity, local autonomy, and so on".[182] Despite differences in opinions and beliefs, otherwise disparate communities within the state nonetheless prove capable of forming an ever-evolving national identity, or at least an active, ongoing partnership, on the basis of these common understandings. The specific form of the political state itself emerges as the result of these deliberations between smaller communities, as the collective decisions of citizens ultimately shape the institutional structures of which the state is comprised.[183] Insofar as these groups have common ground in the form of shared vocabularies, understandings, and agreements on processes and shared institutions, it is possible for them to co-exist and thrive under the broader umbrella of a sovereign state.

Under Walzer's conception of community, the bridge between civil society and the political state is built upon a foundation of common understanding. This common understanding is both reflected in, and itself influenced by, a community's culture. Culture often serves to reaffirm values, both moral and political, deemed by a community to be positive, while denouncing those perceived to be detrimental. Accordingly, culture serves an integral role in the development of the common understanding necessary for meaningful cooperation and self-determination. Whereas in the past exposure to culture would be limited due to it being conveyed predominantly through shared experiences rendered possible by geographic proximity, the advent of communication technologies has dramatically shifted the way culture is delivered. The introduction of the radio and television has served in part to strengthen the basis of common understanding, as both state-funded and public media have offered more immediate access to information relevant to both the state and the cultural communities of which it is comprised. Similarly, the later introduction of the internet has offered further avenues through which culture can be immediately delivered and experienced. Through shared exposure to culture, otherwise

[181] Walzer, *Spheres of Justice*, 29.
[182] Walzer, *Spheres of Justice*, 29.
[183] Walzer, *Spheres of Justice*, 29.

disparate communities strengthen their common understanding which, in turn, contributes to the formation of a political state which best reflects their shared values.

It is this common understanding which sets the stage for cooperation amongst sub-communities and makes the common life possible. This common life serves as a prerequisite for the unity of the state, or at least its unified integrity in acting for and representing some of the most important interests of its members, like their physical security and their desire to make their own most vital life-decisions. As a result, it is not only the immediate and particular lives and well-being of its citizens that the state is tasked with protecting, but also the common life from which the state derives its societal values and norms, and around which the unique identity of the state takes shape. The extent to which this common life is cherished by the citizens of the state determines the justifiability of a state's efforts to defend it against malign foreign intrusions; states which protect a common life that defends individual liberties and which is seen by the citizenry as worth incurring sacrifices to upkeep are seen as having greater moral standing than those states which merely preside over a "common life" of subjugation and oppression.[184] Accordingly, if a state's common life is absent, or the state makes no effort to defend a cherished existing common life, its own defense lacks moral justification. Ultimately, the preservation of this common life motivates the rights of territorial integrity and political sovereignty, both of which serve to defend it against the machinations of outsiders.

These baseline assertions regarding the common life underpin what Walzer refers to as "the legalist paradigm" of aggression. The legalist paradigm can be broken into six main propositions, the first two of which are concerned with the overall existence of an international society of independent states and the corresponding law they adhere to, granting them the rights to territorial integrity and political sovereignty.[185] From these first two propositions, the legalist paradigm asserts that uses of force employed against the rights of another state ultimately constitute criminal acts of aggression which, in turn, justify two specific types of forceful responses: wars of self-defense and wars of law enforcement.[186] The first of these types of wars

[184] Walzer, *Just and Unjust Wars*, 54.
[185] Walzer, *Just and Unjust Wars*, 61.
Proposition 1: "There exists an international society of independent states."
Proposition 2: "This international society has a law that establishes the rights of its members – above all, the rights of territorial integrity and political sovereignty."
[186] Walzer, *Just and Unjust Wars*, 62.

represent unilateral responses by a targeted state against their belligerent counterpart, while the latter more resemble coalition efforts against rogue states within which an international body comes to a multilateral consensus regarding the need for intervention, such as the sizable international coalition arrayed against Saddam Hussein's Iraq in the Persian Gulf War of 1991.[187] The legalist paradigm's just cause criterion is established in proposition five, asserting that only aggression stands as just grounds for going to war, while proposition six declares that belligerent states can be punished following military intervention.[188] Collectively, the legalist paradigm offers a foundation upon which the crime of aggression is defined and conceptualized as a morally justified reason for going to war.

Notably, Walzer's analysis of the third proposition of the legalist paradigm contains a key caveat. In seeking to dispel ambiguity regarding the determination of when an act of aggression has occurred, Walzer opts to narrow potential acts of aggression to solely armed invasions and physical assaults.[189] While this narrower conception of aggressive action identifies two specific circumstances within which states are uncontroversially justified in responding in self-defense, limiting acts of aggression to these two instances proves problematic when accounting for cyberwar. As discussed in the previous chapter, the vast majority of cyberattacks fail to result in kinetic-equivalent harms. Most cyberattacks cause intangible harms which fail to manifest as direct physical damage. Likewise, while a "digital invasion" may be theoretically possible, perhaps by an aggressor remotely seizing control of a state's domestic cyber infrastructure with no intention of relinquishing it, it nonetheless seems materially different to invasions conducted through the mobilization of military force across territorial borders and physical occupation. Under Walzer's conception of aggressive acts, disanalogous cyberattacks appear to fall between the cracks, failing ever to qualify as acts of aggression sufficient for motivating a state's inherent right to self-defense.

Proposition 3: "Any use of force or imminent threat of force by one state against the political sovereignty or territorial integrity of another constitutes aggression and is a criminal act."
Proposition 4: "Aggression justifies two kinds of violent response: a war of self-defense by the victim and a war of law enforcement by the victim and any other member of international society."

[187] Alex Danchev and Dan Keohane, "Introduction: The Rules of Propriety," in *International Perspectives on the Gulf Conflict, 1990-91*, eds. Alex Danchev & Dan Keohane. (New York: St. Martin's Press in association with St. Antony's College, Oxford, 1994). xiv.

[188] Walzer, *Just and Unjust Wars*, 62.
Proposition 5: "Nothing but aggression can justify war."
Proposition 6: "Once the aggressor state has been militarily repulsed, it can also be punished."

[189] Walzer, *Just and Unjust Wars*, 62.

I argue that limiting potential acts of aggression to merely these two kinds of offenses is unsatisfactory for the evaluation of cyber operations, and further note that Walzer's landmark contributions were first crafted in the 1970s, long before the Internet revolution. While territorial invasions and physical assaults are readily recognizable as violations of the rights to territorial integrity and political sovereignty, **I argue that cyberattacks prove to be equally capable of posing existential threats to the common life from which such rights are derived**. Despite the majority of cyberattacks failing to bridge the physical-digital divide, such attacks may nonetheless prove harmful insofar as they may target the common life directly in a manner which avoids the same evaluative scrutiny and severity of response which would be incurred by conventional military action. To this end, the subsequent section considers Whetham's work pertaining to tactics within medieval warfare and their evolution within the realm of cyberwar, in order to show that non-physical cyberattacks can be considered acts of aggression capable of grounding claims to self-defense, even in the absence of accompanying conventional action.

4.3 War for the 21st Century: Whetham's *Chevauchées*

The lack of a physical component to cyberwar has led some to decry cyberwar's potential to truly constitute war. For some critics, such as Rid, cyberwar itself cannot happen due to the absence of violence and the presumed inability of purely digital attacks to present realistic threats of violence.[190] This approach is rebuffed by Whetham, who looks towards medieval history and the actions of *chevauchées* to illustrate that the concept of war is not limited to the clashing of combatants meeting head-on; rather, it encompasses a myriad of direct and indirect tactics seeking to undermine an enemy's willingness and ability to resist aggression. For Whetham, the position of those who criticize cyberwar's potential to constitute genuine war closely resembles the dismissiveness with which medieval historians have often treated the actions of *chevauchées*, seeing them as merely occurrences incidental to war, rather than playing a key part within hostilities.[191]

---

[190] David Whetham, "Cyber *Chevauchées*," In *Binary Bullets: The Ethics of Cyberwarfare*, eds. Fritz Allhoff, Adam Henschke, & Bradley Jay Strawser (New York, NY: Oxford University Press, 2016), 75-76.
Thomas Rid, "Cyber War Will Not Take Place," *Journal of Strategic Studies* 35, no. 1 (February 2012): 5–32, https://doi.org/10.1080/01402390.2011.608939. 29.
[191] Whetham, "Cyber *Chevauchées*," 75.

Distinct from armoured formations of soldiers tasked with laying siege to enemy battlements, medieval *chevauchées* were processions of soldiers whose primary objective was the plundering of enemy territory in between pitched battles.[192] Whetham notes that medieval scholarship until the 1950s had a relatively straightforward conception of medieval warfare tactics, viewing direct battles between warring factions as being the real focus of war, whereas the actions of *chevauchées* were seen as an interim pastime, something that was done by restless troops waiting for orders in between battles.[193] Insofar as the primary objective of war was "the destruction or overthrow of the enemy's forces in order to impose one's will on [them]", the *chevauchées* were seen as failing to contribute directly towards the fulfillment of this end. Accordingly, despite the prevalence of *chevauchées* in the conduct of medieval conflict, their actions were largely dismissed by earlier medieval historians as merely an aside to medieval conflict, ultimately irrelevant to military strategy and the fate of states.[194]

The *chevauchées* would eventually be revisited by later medieval scholars who would reconceptualize their role within the context of medieval warfare and ascribe them much greater importance. Rather than merely bored soldiers or opportunistic raiders, *chevauchées* came to be seen as a key strategy of attrition which worked along two dimensions. Firstly, by plundering land adjacent to military strongholds, the *chevauchées* served to destabilize the economic resources of their adversaries; this would, in turn, limit their enemy's ability to endure a protracted military conflict.[195] Secondly, *chevauchées* had actually proven effective at undermining the legitimacy of rivalling sovereigns in the eyes of their own citizens, reducing their subjects' willingness to fight.[196] Whetham states that a sovereign's ability to defend their own territory and people against foreign aggressors determined, at the time, whether it was perceived that the sovereign truly had the right to rule.[197] After all, why obey a state incapable of protecting you? Accordingly, the specter of *chevauchées* marauding unopposed through a sovereign's territory served to undermine a sovereign's position as the rightful ruler, potentially rendering their defensive position untenable and forcing them to the negotiating table with less leverage. The appeal of *chevauchées* is bolstered by the comparative risks inherent to more

[192] Whetham, "Cyber *Chevauchées*," 75.
[193] Whetham, "Cyber *Chevauchées*," 75.
[194] Whetham, "Cyber *Chevauchées*," 75.
[195] Whetham, "Cyber *Chevauchées*," 81.
[196] Whetham, "Cyber *Chevauchées*," 81-82.
[197] Whetham, "Cyber *Chevauchées*," 81.

conventional tactics of the time, as medieval historians note that open battle losses could not only prove costly in terms of wealth and manpower, but also devastating for the morale of either belligerent as the results of direct conflict were often taken by observers as "a clear judgment by God on the merits of their respective claims".[198] Insofar as *chevauchées* proved capable of weakening an adversary's overall ability to fight without the corresponding levels of risk associated with direct battle, *chevauchées* represented more than merely wartime extracurriculars; rather, they themselves constituted acts of war, and quite potent ones at that.

While the specific tactics used differ dramatically, Whetham suggests that modern cyberattacks are heavily reminiscent of the actions of medieval *chevauchées*. Whereas medieval combatants were reliant on physical force to seize a state's wealth or to hamstring its ability to organize a concerted defense, Whetham's "cyber *chevauchées*" prove capable of achieving these objectives through cyber weapons and tactics. To illustrate, Whetham proposes a hypothetical situation between two militarily equal states, hereafter termed State A and State B, locked in a territorial dispute regarding a resource-rich territory. Supposing that State A unilaterally moves into the disputed territory, State B would be forced into a response to dissuade State A from their course of action. While outright conflict may prove undesirable insofar as it may be financially costly and politically unpopular domestically, Whetham posits that State B may instead express their displeasure with a series of unattributed cyberattacks disrupting State A's air defense grid at key points. These cyberattacks would be preceded by anonymous public announcements declaring which element of the network was to be shut down, and subsequently followed by State B disavowing their role in the matter. As such attacks continue, State A would be publicly embarrassed and held to account by citizens made to feel "fearful, vulnerable, and unprotected" against the will of a foreign actor.[199] Whetham suggests that the undermining effect of State B's pervasive cyberattacks, coupled with State A's inability to stop or deter the attacks, would likely lead to State A offering concessions to State B in exchange for the cessation of their cyber operations.[200] Accordingly, State B will achieve their foreign policy interests without resorting to the mobilization of military force.

[198] Whetham, "Cyber *Chevauchées*," 81.
[199] Whetham, "Cyber *Chevauchées*," 82-83.
[200] Whetham, "Cyber *Chevauchées*," 83.

In this example, the cyber offensive launched by State B heavily resembles the actions of medieval *chevauchées*, working to shake the foundation of legitimacy enjoyed by a foreign state in hopes of advancing a foreign policy objective. Evidently, the character of medieval *chevauchées* proves replicable within cyberspace, as the development of cyber weapons and tactics offers states a comprehensive toolkit with which to approach interstate disputes. States are now capable of remotely targeting the cyber infrastructure of others to hit a myriad of targets, from military networks to civilian data, and to exert their political will abroad. Much like their medieval counterparts, cyber *chevauchées* could be politically destabilizing, as evidenced by Russia's efforts during their 2014 Crimean campaign within which targeted cyberattacks sought to foment irredentist aspirations amongst a Crimean populace disillusioned with the Kyiv-based Ukrainian government. Beyond being a particularly effective tool for achieving foreign policy goals, cyberoperations also offer a level of plausible deniability which makes otherwise difficult diplomatic positions more manageable for aggressive states. Whetham acknowledges that, in cases such as the hypothetical one posed above, State A would likely suspect State B's responsibility for the cyberattacks plaguing their sphere of influence, however he adds that cyber belligerents are likely to adopt a "correlation is not causation" approach to the situation; State B would remain stern in its denial of responsibility for the ongoing attacks, while tacitly suggesting that such cyber offensives may be halted should State A cease its operations within the disputed territory or make concessions towards State B.[201] This sort of rhetoric was on display during the 2007 cyberattacks on Estonia's digital infrastructure, within which Russia's government consistently denied responsibility while simultaneously suggesting that the attacks may have been the work of pro-Russian independents upset with Estonia's political decisions.

The actions of cyber *chevauchées*, both in Whetham's scenario and beyond, are unlikely to be regarded as acts of aggression under current evaluative frameworks. The absence of a physical damage- or bodily harm component resulting from cyber *chevauchées* ensures that such acts would fail to qualify as armed attacks necessary for fulfilling the just cause criterion under the analogue-reliant framework offered by the *Tallinn Manual*. Such cyberattacks similarly fail to constitute either territorial invasions or physical assaults which Walzer denotes as necessary for asserting that an act of aggression has occurred between states. A cyberattack is unlikely to

[201] Whetham, "Cyber *Chevauchées*," 85.

be considered tantamount to a physical assault unless it results in the sort of injury or destruction typically wrought by kinetic attacks. Likewise, purely digital attacks could potentially be argued as constituting a territorial invasion of sorts, however such a declaration would likely prove controversial due to disagreements regarding the meaning of territory and the right of ownership within cyberspace. This would place cyber *chevauchées* as yet another short-of-war activity which, despite possessing the potential to be incredibly disruptive and to lead to diplomatic tensions between states, nonetheless falls short of fulfilling the just cause criterion.

Dismissing cyber *chevauchées* in this manner appears short-sighted given that these kinds of attacks are emerging as not merely an occasional alternative to conventional action, but rather as the new norm within interstate conflict. While conventional attacks uncontroversially present a clear and imminent danger to the common life serving as the foundation for states, so too can cyberattacks prove to be an existential threat. Pervasive cyberoperations can work towards hindering a state's ability to guarantee the rights of its citizens. Attacks against digital infrastructure underpinning the operations of banks and hospitals can manufacture discontent with a state's political governance which would otherwise not have existed barring foreign interference. Likewise, continued digital attacks against a state's defensive capabilities such as that proposed by Whetham would serve to foment the same sort of discord which could exert undue strain on a state's citizens who find themselves at the mercy of a foreign aggressor, without any of the defense that membership of a state ought to provide. Insofar as such attacks fail to cross the thresholds identified by Walzer and the *Tallinn Manual*, these cyber operations ensure a slower and less severe response than a conventional attack would garner, affording cyber operations a greater overall impact than would be reasonably achievable through regular means, without the associated drawbacks.

Given the rising frequency of non-destructive cyberattacks accompanying the paradigm shift of conflict away from the conventional domains and towards cyberspace, cyberattacks may prove a greater existential threat to the common life of 21st century states than kinetic alternatives. This, in turn, suggests that the majority of acts of aggression within cyberspace will not manifest as cyberattacks resulting in explosions of pipelines or the remote detonation of stockpiled munitions as some have come to fear. Rather, the target of many cyberattacks is instead the common life of a state, as attacks against the interests and rights of a populace serve to throw states into disarray and render them either more amicable to concessions or less capable

of mounting a concerted defense against further aggression. Insofar as current frameworks decline to consider such non-destructive digital attacks as acts of aggression, targeted states find themselves in a bind as they remain vulnerable to the cyberattacks of a foreign belligerent, without being able to respond with any degree of force in turn.

As a result, it is necessary to rethink our conception of when the just war condition of self-defense has been met. Despite the lack of territorial invasion or physical assaults, certain cyber offensives prove every bit as violative of a state's common life as their kinetic counterparts. However, as evidenced in the preceding chapter, cyberattacks failing to result in outright destruction are often perceived as failing to cross the threshold of just cause for self-defense specifically due to the absence of these kinds of harms. Likewise, in the legalist paradigm account for aggression, the lack of physical assaults or territorial invasions by cyber means limits the likelihood that a cyber operation could ever be seen as an act of aggression satisfying the just cause criterion which triggers the right to self-defense. This is problematic insofar as the common life, the critical element of a state whose well-being motivates our traditional understanding of the just cause criterion, is vulnerable to suffering immense damage resulting from disanalogous cyber offensives. Accordingly, there is an apparent need to adopt a novel account for the evaluation of cyber aggression to encompass the harm caused by non-physical cyberattacks and to offer an avenue for states to designate some such attacks as acts of aggression satisfying the just cause criterion, either in the interests of motivating immediate self-defense, or for the escalation of such issues to international forums such as the UN.

4.4 Smith's Sovereignty Account of Just Cause

One such attempt at addressing the turbulent grey area of just cause existing between conventional military action and non-destructive cyberattacks is offered by Smith as a response to the current status quo approach to extending traditional understandings of aggression into cyberspace. Smith takes exception to what he terms the "Standard View" – the belief that cyberattacks have the potential to fulfill the just cause criterion if and only if they produce

physical effects.[202] The Standard View echoes the sentiments expressed within the *Tallinn Manual*, asserting that cyberattacks stand as reasonable grounds for invoking the right to self-defense should their "scale and effects" be akin to those of conventional attacks. While the Standard View does well to distance itself from the overly myopic conception of aggression posited by a traditional approach which wholly excludes cyber operations, the shift to a consequence-based model of evaluation raises further concerns for Smith, who notes that a variety of non-violent means, such as trade embargos and sanctions, often result in physical effects equivalent to kinetic military strikes; for example, widespread embargos could result in physical destruction via ensuring that degrading state infrastructure falls into disrepair without the ability to source vital resources from abroad.[203] While the end result of infrastructure damage is the same, we would be hard-pressed to claim that embargos and kinetic strikes lack a material difference. Evidently, there must exist some factor beyond mere consequences to distinguish an act of aggression from acts which may inadvertently result in physical damage or bodily harm.

In response to the hegemonic Standard View, Smith proposes an alternative he coins the "Sovereignty View". Instead of placing emphasis on the physical consequences of aggressive actions, the Sovereignty View focuses on the *methodology* and *effect* of attacks on states' "collective, political self-determination".[204] In the vein of Walzer's earlier work on the rights which serve as the foundation of a state, Smith points to two specific rights he posits to be integral to a state's ability to practice self-determination. The first of these is the right to political autonomy, namely the ability of collectives to self-determine their own values and the policies most conducive to upholding them.[205] The second right is that of territorial integrity, as Smith notes that the successful implementation of self-determined policy is only made possible by holding territory within which such policies may be enacted.[206] Taken in tandem, these two rights form the foundation of a state's sovereignty, much as with Walzer. Accordingly, the Sovereignty View's evaluation of just cause asserts that the just cause criterion is satisfied when

[202] Patrick Taylor Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account: Cyberattacks as Casus Belli," *Journal of Applied Philosophy* 35, no. 2 (May 2018): 222–41, https://doi.org/10.1111/japp.12169. 222.
[203] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 225-226.
[204] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 227.
[205] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 227.
[206] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 227.

a foreign state violates either of these two rights, irrespective of whether such an action is conducted by conventional or digital means.

While deliberations regarding whether cyberattacks can violate territorial integrity remain unresolved, it is more readily apparent that cyberattacks have the potential to violate the political autonomy necessary for political self-determination. With regards to states heavily reliant on digital infrastructure for their domestic elections, it is well within the realm of possibility for foreign operatives to penetrate potential computer defenses and manipulate the results to reflect a desired outcome, such as the election of a head of state with interests more aligned with those of the interfering entity. Beyond cyberattacks conducted directly against the state, Smith suggests that attacks against non-state institutions may likewise comprise violations of political self-determination should they "[attempt] to rearrange, undermine, or violate the legal and political entitlements that we have collectively decided will be the policy within our borders".[207] To illustrate, Smith proposes a hypothetical situation within which State A seeks to influence port policies within State B to facilitate an increase in their own relative economic strength. One option at State A's disposal is the implementation of sanctions meant to hurt State B's interests and motivate it towards adopting a desired policy change. Alternatively, rather than targeting the state directly, State A could instead "covertly enter the stevedore union offices and change electoral results so that the candidate favourable to their interests is elected".[208] The election of a preferred union candidate could serve to have State B's port unions adopt policies conducive towards State A's overall objectives, while maintaining the appearance that the election of the union representative and the subsequent adoption of their policy interests have occurred organically, without foreign interference, and that both are an accurate reflection of the self-deliberation of the relevant parties.

While this latter option is not a direct attack against the state itself, Smith nonetheless asserts that it represents a violation of the political self-determination of the citizens of State B on the grounds that the citizens have "collectively determined that unions have a particular legal structure, elect officers in particular ways (or have legally determined discretion concerning their leadership structure) and these determinations have economic and political consequences that

[207] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 228.
[208] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 228.

[State A] is attempting to undo".[209] This coincides with Walzer's notion of the common life, insofar as he views societal institutions as the direct products of sincere deliberations amongst the component communities which comprise the state. Under the Sovereignty View, these kinds of operations could qualify as just cause for military action on the grounds that they are ultimately violations of State B's political sovereignty. It is worth quickly adding here that this would only serve to motivate the just cause criterion which alone is insufficient for the actual mobilization towards war. The remaining *jus ad bellum* principles likewise need to be satisfied prior to morally justifying the decision to go to war.

Critics may suggest that the Sovereignty View falls victim to the same pitfall which plagued the *Tallinn Manual* deliberation process and ultimately led to their endorsement of the Standard View. As Smith notes, there is no shortage of cyberattacks which, while violative of state sovereignty, nonetheless seem to be regarded as generally acceptable amongst the global community.[210] Cyber espionage, for example, is tacitly accepted by the international community. DDoS attacks, such as those launched against Estonia, are seen as disruptive, but ultimately falling short of demanding an immediate military response. The shift to the Sovereignty View requires establishing thresholds or criteria for distinguishing cyberattacks which violate a state's right to political autonomy from those which fail to pose an imminent threat to a state's ability to practice political self-determination. There must be a standard for separating milder cyberattacks which violate sovereignty but do not hinder political autonomy, such as temporarily disruptive DDoS attacks, from the more severe cases within which political self-determination is undermined or rendered impossible, such as an attack on a sovereign state's electoral processes.

Smith's response to these concerns comes in two parts. Firstly, Smith asserts that a cyberattack which imposes costs sufficient for restricting a state's ability to practice political self-determination satisfies the just cause criterion.[211] Secondly, and more importantly, Smith suggests cyberattacks constituting a direct imposition of will by one state on another similarly motivate just cause. This distinction is contingent on "whether the unfriendly strategic action aims to achieve its political objective *through* or *in spite of* the deliberations of the target

[209] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 228.
[210] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 228.
[211] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 231.

state".[212] States are free to operate in a manner designed to motivate a target state towards a specific politically relevant outcome, such as the adoption of friendly foreign policy, however they are prohibited from manipulating their political deliberations directly, and in coercive and under-handed ways. In line with Smith's earlier hypothetical, State A's sanctions designed to encourage capitulation by State B may impose a cost on State B, but nonetheless fail to qualify as just cause. While such sanctions may strain State B's interests, they are merely one factor to be accounted for within their otherwise unconstrained political deliberations; State B remains free to choose otherwise.[213] This would be an example of State A seeking to achieve foreign policy objectives *through* the deliberations of a target state. By contrast, the alternative approach undermining the self-deliberations of State B's citizens represents a cyberattack which achieves its objective *in spite of* State B's deliberations; State B is incapable of choosing its course of action for itself because of State A's interference. While both approaches may result in a change in State B's port policy, the former approach is permissible insofar as it preserves State B's right to political self-determination. By contrast, the latter is impermissible insofar as it violates that right.

Under Smith's Sovereignty View, cyberattacks constitute just cause for unilateral military response if they serve to "overwhelm, bypass, or override resistance by the target state" and do so to achieve a specific politically relevant objective.[214] Accordingly, Smith asserts that a wide range of cyberattacks inherently fail to constitute just cause. Cyber espionage operations commonly seek to avoid resistance entirely, and their objective is typically the acquisition of sensitive information rather than directly shaping a foreign state's political deliberations.[215] Smith further argues that the vast majority of DDoS attacks likewise fail to satisfy the just cause criterion on account of their methodology being relatively unintrusive, insofar as they simply aim to overwhelm digital infrastructure from the outside rather than directly penetrating a state's networks, and seem fairly simple to respond to.[216] As a result, the DDoS attacks on Estonia in the wake of the controversial decision to relocate the Bronze Soldier of Tallinn fail to satisfy the just cause criterion under the Sovereignty View. While the attacks undoubtedly imposed a cost on

[212] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 228.
[213] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 228.
[214] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 231.
[215] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 233.
[216] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 233.

Estonia in hopes of changing the Estonian government's stance, the attacks failed to constitute an imposition of will; the Estonian government remained capable of political self-determination even amidst the ongoing attacks. It is worth noting that Smith acknowledges that a severe DDoS attack may theoretically fulfill the just cause criterion, however not due to it constituting an imposition of will. Rather, a DDoS attack meets this criterion should it manage to have a severe effect on a state's ability to practice self-determination, such as through shutting down cyber infrastructure vital to a government's ability to govern and fulfill the functions of a state for a prolonged period.[217]

By contrast, the Sovereignty View considers penetrative cyberattacks whose methodology seeks to overcome resistance and result in "cyberharm" to be tantamount to an armed attack for the purposes of just cause.[218] Smith suggests that the combination of legal and political entitlements to the undisrupted functioning of a crucial computer system (namely, those integral to a state's ability to govern and self-deliberate, such as electoral systems), coupled with efforts made to defend such systems from foreign interference, render intrusive attacks unjustified impositions of will.[219] Consequently, cyberattacks such as Stuxnet satisfy the just cause condition insofar as the Stuxnet worm was specifically designed to undermine Iranian resistance in the form of numerous security measures designed to protect their nuclear infrastructure from being assailed by foreign states; Smith asserts that the development and subsequent deployment of Stuxnet served to "impose a particular outcome" upon Iran, one that wholly disregarded and, in fact, overrode Iranian deliberation.[220] Cyberattacks need not result in physical consequences in order to qualify as just cause. Cyber operations seeking to penetrate the defenses of electoral software in order to manipulate the results would likewise satisfy the just cause criterion, even in the absence of direct physical consequences. Notably, this is not limited by scale, as the target need not necessarily be national elections; in Smith's earlier hypothetical scenario cyberattacks against the stevedore union offices of State B were considered sufficient for constituting an imposition of will by virtue of ignoring the self-deliberation of its citizens. Despite the target of the attacks being of a lower profile, the intrusive nature of the attack and the

[217] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 233.
[218] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 233.
"[Cyberharm] is intentional harm by an agent, *via* an infomatics network such as the Internet, in which the functioning of a system (a person, a machine, software, or an economy) is in some way impaired or degraded."
[219] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 233-234.
[220] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 234.

ultimately political nature of its objective render it sufficient for fulfilling the just cause criterion under the Sovereignty View.

The appeal of the Sovereignty View is twofold. Firstly, the Sovereignty View is better equipped to account for a wider spectrum of cyberattacks, rather than solely the fringe cases which may result in harms analogous to conventional action. By rescinding the scale and effects criteria from just cause calculations, the Sovereignty View offers a pragmatic avenue towards the evaluation of the kinds of non-physical cyberattacks which have become the norm of interstate conflict in the 21st century. Importantly, this approach renders it likely that a broader range of cyberattacks potentially qualify as just cause, unlike under the Standard View in which only the most extreme and unlikely kinds of attacks potentially satisfy the criterion. Loosening the just cause criterion here would offer states greater flexibility in responding to cyberattacks, as the just cause criterion may motivate stronger defensive measures, even if not escalating to the deployment of force. Likewise, acknowledging that a wider range of cyberattacks may feasibly constitute just cause could encourage greater restraint on the part of belligerents, unlike in the current status quo within which the vast majority of cyberattacks, save those causing death or destruction, are seen as fair game.

Secondly, the Sovereignty View's emphasis on the protection of state rights to political self-determination better serves to preserve the common life, which is the motivational force to the state's inherent right to self defense, than the Standard View does. While Walzer conceives of just cause as manifesting in the form of territorial invasions and physical assaults, it is no longer the case that only these types of hostile action pose a threat to this common life. Instead, foreign actors find themselves capable of assailing the common life, and the common understanding on which it is predicated, through purely digital means, avoiding engaging in unnecessary physical armed conflict in pursuit of their foreign policy objectives. Under the Standard View, non-physically harmful attacks are largely dismissed as falling short of just cause due to their failure to result in scale and effects akin to "real" armed attacks. This results in a myriad of cyberattacks posing an existential threat to the common life, such as widespread penetrative interference in democratic processes, falling short of fulfilling the just cause criterion designed to afford states the right to defend against such threats. By contrast, the Sovereignty View loosens the physical effects prerequisite of the Standard View to accommodate the wider range of cyberattacks which not only present similar, or perhaps graver, threats to the common

life but are also those which are more likely to threaten states within the 21st century. By rendering it likelier that a cyberattack may constitute just cause, the Sovereignty View both: 1) enables targeted states to better defend themselves against incoming cyberattacks; while 2) encouraging belligerents to rethink their cyber operations now that fewer types of cyberattacks could be conducted with impunity.

Smith acknowledges that the Sovereignty View's move to broaden the range of cyber operations capable of fulfilling the just cause criterion may face charges of being overly permissive, rendering even relatively trivial cyberattacks potential just cause insofar as they represent an imposition of will.[221] The foreign intelligence apparatus of State A may respond to State B's vote to ban foreign ownership of businesses within its borders with concerted cyberattacks against its Parliament, seeking to change the submitted ballots in order to prevent the ban from going through. While such cyber operations undoubtedly represent an imposition of will by State A on State B, it is counterintuitive (to some) to claim that a full-fledged military counteroffensive represents a reasonable response. Smith suggests these concerns are ultimately immaterial, as just cause alone is insufficient for motivating the deployment of military force in the interests of self-defense; Smith notes that, in cases within which the violation of sovereignty is ultimately considered trivial, the proportionality and necessity criteria of JWT would see to it that forceful responses are nonetheless prohibited.[222] The looser just cause threshold introduced by the Sovereignty View remains tempered by the constraints put in place by subsequent JWT principles.

As a result, the Sovereignty View offers a stronger evaluative means of determining whether the just cause criterion has been met by any given cyber operation than that provided by the Standard View. By diminishing the importance of physical harm as a consequence, the Sovereignty View is better equipped for the evaluation of a fuller spectrum of cyberattacks which a state may reasonably face, rather than the predominantly speculative worst cases such as a cyber-Pearl Harbor. Likewise, the Sovereignty View better reflects a commitment to the preservation of the values which inspire states' rights to self defense than the Standard View does. Rather than downplaying the potential of non-physically harmful cyberattacks as grounds for going to war, the Sovereignty View acknowledges that such operations may pose an

[221] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 235-236.
[222] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 236.

existential threat to a state's common life and, by extent, its ability to practice political self-determination. Insofar as the moral justification of a state's defensive action is contingent on its defending of a cherished common life, the Sovereignty View proves better aligned with the intent of the right of self-defense. Provided a foreign state deploys cyber operations to impose its will in spite of the deliberations of the target state, the just cause criterion is satisfied for the latter.

4.5 Beyond Smith: Evaluating Unjust Cyber Interference

While the Sovereignty View does well to incorporate the threat cyberattacks pose to political self-determination into discussions of just cause, I argue that Smith's account should nonetheless be extended to further encompass yet another type of cyber operation which works along a different axis. In its current form, the Sovereignty View identifies two specific types of cyberattacks distinguished by virtue of how they achieve their objectives. The first of these are cyberattacks which seek to shape foreign policy through a state's own organic deliberation. These types of attacks may be disruptive or impose a cost on a state to motivate them towards a specific political outcome, much like the imposition of sanctions may work towards encouraging states to adopt favourable policies. In these cases, the final decision ultimately remains within the sole power of the targeted state. By contrast, the second type of cyberattack specifically aims to undermine a state's deliberations in order to bring about a particular result. These attacks work to subvert a state's ability to practice political self-determination. These latter types of attacks are more direct than the former, overriding a state's deliberations entirely rather than offering an impetus to change the course of their deliberations themselves. Under the Sovereignty View, the first type of cyberattack falls short of just cause insofar as the targeted state is still capable of practicing political self-determination. By virtue of undermining this capability, the second type of attack fulfills the just cause criterion.

I argue that the Sovereignty View's dichotomy of cyber operations working *through* state deliberation and working *in spite of* state deliberation fails to properly account for a third type of cyber operation which occupies the grey area between the two. This third type of cyber operation neither seeks to impose direct costs nor directly override the deliberations of a state. Rather, this

third type seeks to *manipulate* the deliberations directly, corrupting the common life of a state so that its deliberations align with the aggressor's political objectives, while seemingly appearing to be the uncoerced will of the target state and the citizens it represents. These cyber operations often take the form of coordinated, systematic disinformation campaigns which seek to aggravate political discontent within foreign states, broadening divides between groups and sowing mistrust or resentment towards governments which have proven reluctant to bend to an aggressor's interests. The result is an *inorganic* common life which has been deliberately manipulated by a malign foreign aggressor in accordance with its specific political objectives.

Whereas the two kinds of operations encompassed within the Sovereignty View target the state, this third type targets civil society directly. Most states harbour some degree of internal division along cultural, ethnic, political, or religious lines. However, in most cases, these divisions are bridged by a shared common understanding which motivates cooperation and self-deliberation within a political whole. This third kind of operation seeks to accomplish foreign policy objectives through distorting this common understanding and/or inorganically exacerbating these divisions, either to promote a specific political outcome or to undermine the state's general ability to self-deliberate. To this end, these operations may disseminate disinformation to remove this common ground and instead further stratify these sub-communities, rendering them more likely to believe the values of the communities are, at best, irreconcilable, or, at worst, fundamentally opposed.

This third kind of cyber operation is rendered more effective by virtue of how contemporary culture is delivered. Much culture is presently experienced through digital mediums, be it television, radio, or the internet. As such, a state's culture is rendered vulnerable to foreign intrusions. Traditional sources of news and culture, such as radio and television, have been targeted by cyberattacks, with cyber operatives proving capable of taking down systems necessary for conveying information, such as digital video libraries.[223] The rise of social media and internet forums has offered cyber operatives further avenues of shaping public opinion. While social media may theoretically serve as a forum within which the organic culture of various sub-communities comes to the forefront, it presents a tantalizing target for cyber

[223] Jake Tapper, Evan Perez, and Ryan Young, "Cox Media Group hit by cyberattack last week, sources familiar tell CNN," *CNN*, June 9, 2021. Accessed January 16, 2022. https://www.cnn.com/2021/06/09/politics/cox-media-group-cyberattack/index.html.

operatives who wish to shape community discourse or sow division between communities. Domestically, social media cyberoperations work "(1) to suppress fundamental human rights; (2) to discredit public opposition; and (3) to drown out political dissent".[224] These operations may similarly be deployed as a component of foreign influence campaigns working towards shaping public opinion towards the acts of a foreign state, as evidenced by China's social media operations in response to the 2019 anti-extradition law amendment protests in Hong Kong, during which the Chinese government used various social media platforms to "paint Hong Kong's democracy advocates as violent radicals with no popular appeal".[225] In both domestic and foreign contexts, there is a concerted effort to shape the common understanding at the heart of the state towards a politically desirable outcome, be it the widespread adoption of a particular opinion, the stifling of dissent, or the magnification of disagreements so as to undermine the likelihood of cooperation amongst diverse sub-communities, be they political, cultural, religious, or otherwise.

The effect of these kinds of operations is perhaps made most starkly apparent by the phenomenon of "fake news". The manipulation of politically charged news serves not only to evoke reactions within a targeted community, but also to undermine common ground by generating uncertainty and disagreement regarding basic facts and values. Throughout the 2016 presidential election cycle and beyond, "fake news" has emerged as a politically charged phrase used to deride the opinions of others as being fundamentally misinformed. In a 2018 study asking US adults to define "fake news", Tong et al. found that 34.2% of responses failed to define the term neutrally, with the definitions provided instead "incriminat[ing] opposing political and media entities for the predicament of fake news".[226] Insofar as the veracity of basic facts can be rendered controversial through mass disinformation, finding common understanding between political groups becomes a more arduous task. In these cases, the divide between political sub-communities is no longer regarded by the sub-communities themselves as merely a difference of opinion, but instead the result of the "other" community being fundamentally

[224] Samantha Bradshaw and Philip N. Howard, "The Global Disinformation Order: 2019 Global Inventory of Organised Social Media Manipulation," *Oxford Internet Institute,* (2019). Copyright, Fair Use, Scholarly Communication, etc. 207. https://digitalcommons.unl.edu/scholcom/207. 2.
[225] Bradshaw and Howard, "The Global Disinformation Order", 2.
[226] Chau Tong et al., "'Fake News Is Anything They Say!' — Conceptualization and Weaponization of Fake News among the American Public," *Mass Communication and Society* 23, no. 5 (September 2, 2020): 755–78, https://doi.org/10.1080/15205436.2020.1789661. 769.

disassociated from reality. Understandably, this rift renders cooperation between sub-communities much more unlikely, especially in a democracy which requires some sincere goodwill and consensus to pass majority-based legislation.

These types of operations represent a staple in the Russian hybrid warfare playbook, with many contemporary Russian cyberoperations having been launched with the direct intent to "cause and feed instability, to weaken the social fabric within a society and to complicate and undermine decision-making".[227] Russia's annexation of the Crimean Peninsula was itself predicated by such efforts. As Kyiv underwent political turmoil following the departure of the pro-Russian former president Yanukovych, Russian-affiliated cyber operatives sought to further the divide between Crimean citizens and the distant Ukrainian government in Kyiv. The resulting disinformation campaign worked towards fostering discontent with the national government amongst the predominantly ethnically Russian Crimean populace through misrepresenting the events in the capital as the product of Western-backed nationalists diametrically opposed to the values of ethnic Russians residing within Ukraine. To this effect, misrepresented photos were circulated purporting to show refugees fleeing towards the Polish border en masse, while websites impersonating pro-Ukrainian news sources were brought online to further push the narrative that the new Ukrainian government posed an existential threat to the way of life of ethnic Russians.[228] The net effect of these cyber operations was the weakening of the Ukrainian state's horizontal and vertical legitimacy in the eyes of the Crimean people; *horizontal* legitimacy here refers to "the extent to which the population of a state accept their inclusion in [it]" while *vertical* legitimacy concerns "the extent to which the population of a state accept the right to rule of those who rule".[229] The former would be undermined by the narrative that the Ukrainian state had succumbed to the will of pro-Ukrainian nationalists insensitive to the values of ethnic Russians, while the latter would fall victim to mass disinformation regarding the actions and political affiliations of the incoming Ukrainian government. The result was a fracturing of the common life within the Ukrainian state, as existing divisions between regions

[227] Flemming Splidsboel Hansen and Dansk Institut for Internationale Studier, *Russian Hybrid Warfare: A Study of Disinformation*, 2017, http://pure.diis.dk/ws/files/950041/DIIS_RP_2017_6_web.pdf. 10.
[228] Emilio J. Iasiello, "Russia's Improved Information Operations: From Georgia to Crimea." *The US Army War College Quarterly: Parameters* 47, no. 2 (2017): 7.
[229] Flemming, *Russian Hybrid Warfare*, 10.

and ethnicities were sharply exacerbated by foreign interference seeking to accomplish a political objective: namely, the annexation of the Crimean Peninsula by Russia.

A similar campaign was undertaken in advance of the 2016 US presidential election, with Russian cyber operatives conducting a wide-reaching disinformation campaign in hopes of muddying political discourse and amplifying existing divisions amongst US voters. In addition to the selective release of sensitive Democratic National Committee (DNC) emails, Russian operatives created scores of false individual and group social media presences in order to shape the online narrative surrounding the election.[230] The Russian campaign throughout the electoral process contributed to what McKay and Tenove term *corrosive falsehoods*; namely, efforts which "promote misperceptions and undermine sources of higher epistemic quality".[231] By strategically spreading and promoting disinformation across mediums such as social media, cyber operators worked to subvert the epistemic authority of traditionally relied upon sources of information, instead fostering the sentiment that "truth claims, including expert claims, are largely dictated by political commitments".[232] The result is a sort of epistemic free-for-all within which tensions between members of the state are further exacerbated by the inability to find common ground. This once more leads to the disruption of the state's horizontal and vertical legitimacy; the former by escalating the severity of the rifts between disparate groups of citizens within the state into full-blown and enduring animosity, and the latter by casting doubt on everything from electoral results to the motives behind the ruling party's political decision-making. The lingering effects of this operation would last well beyond the election itself, ushering in today's era of radical—perhaps even poisonous—partisanship within US politics.

In both cases, cyber operations were launched, with malice, to serve inorganic—indeed, foreign—political objectives. However, neither case represents an imposition of direct financial cost meant to motivate the target state towards a specific outcome. Likewise, the operations did not seek to penetrate cyber defenses and directly override democratic elections to install a favourable candidate. Rather, in both cases widespread disinformation campaigns were used to corrupt the common life at the heart of the target state in order to bring about a specific

[230] Spencer McKay and Chris Tenove, "Disinformation as a Threat to Deliberative Democracy," *Political Research Quarterly* 74, no. 3 (September 2021): 703–17, https://doi.org/10.1177/1065912920938143. 707.
[231] McKay and Tenove, "Disinformation as a Threat to Deliberative Democracy," 708.
[232] McKay and Tenove, "Disinformation as a Threat to Deliberative Democracy," 708.

politically relevant outcome which the foreign intervenor desired to achieve. In the case of the US election, rampant disinformation served to further stratify the domestic political discourse to weaken the US' influence globally, the result being more relative international freedom for Russia to move and operate. And perhaps more, such as helping incline US voter choices in 2016 away from a presidential candidate seen as more hostile to Russia towards one perceived as more favourable and pliable. To this end, rather than seeking to convince non-partisan voters to vote for the Republican candidate, Russian information operatives instead focused on animating the entrenched voter base towards a high turnout, mobilizing fervent Republican voters by fueling their fears that the "culture championed by the Democrats is out of touch with [their] values".[233] As part of this campaign, Russian disinformation efforts specifically sought to exacerbate tensions and amplify existing rifts between Evangelical voters and members of other communities within the US, most notably Muslims, the black community, and illegal immigrants.[234]

Meanwhile, in Crimea, similar efforts were used to foster anti-Kyiv and pro-Russian sentiment amongst Crimeans, smoothing the way towards Crimean annexation while hindering the central Ukrainian government's ability to respond. The annexation was itself preceded by a lengthier information war designed to "influence the Russian diaspora in Crimea and convince the world that Ukraine, which was previously part of the Soviet Union, is not a state and has no independent culture".[235] To this end, Russian information operators made heavy use of disinformation and propaganda distributed by Russian-backed media in the preamble to the invasion, later supplemented with coordinated seizures of key media infrastructure within the region to mitigate the Ukrainian government's ability to intervene within the information sphere. The combined effort ultimately "altered the identity of Crimeans and solidified their nascent Russian identity", undermining the common ground between Crimeans and the rest of the Ukrainian community, and rendering them more amicable to annexation into a country that they had "been conditioned to believe was theirs all along".[236] Despite failing to fall neatly into either distinction offered by the Sovereignty View, these attacks represent a clear interference into the

[233] Kathleen Hall Jamieson, *Cyberwar: How Russian Hackers and Trolls Helped Elect a President* (New York, NY: Oxford University Press, 2018). 98-99.
[234] Jamieson, *Cyberwar: How Russian Hackers and Trolls Helped Elect a President*, 102.
[235] T. S. Allen and A. J. Moore, "Victory without Casualties: Russia's Information Operations," Parameters 48, no. 1 (2018), https://press.armywarcollege.edu/parameters/vol48/iss1/8. 63.
[236] Allen and Moore, "Victory without Casualties: Russia's Information Operations," 64.

common life of a state, serving to inorganically alter the fabric which makes up the state's unique identity to coincide with the aggressor's foreign policy objectives.

The intuitively problematic nature of these attacks is not lost on the international community. In early 2022, explicitly decrying the threat disinformation attacks pose to "the free formation of opinion", Sweden has announced its intent to create a Psychological Defense Agency tasked with combating the spread of fake news and defending against foreign interference into its society, with France likewise having put forth plans for a similar agency to defend the integrity of its elections.[237] Sweden's decision comes in the wake of not only the Russian interference campaign against the 2016 US election, but also the spate of disinformation circulating throughout Sweden regarding the COVID-19 pandemic and the state's pandemic measures.[238] As cyber operatives continue to develop novel means of undermining state sovereignty and achieving foreign policy objectives, it is likely that many members of the international community will develop similar agencies in the coming years. This is not unwelcome in light of the theory offered here.

Consequently, I argue that cyber operations seeking to achieve political objectives through *manipulation* of a state's common life likewise satisfy the just cause criterion. By seeking to manipulate the common life of a state directly, an aggressor poses an existential threat not only to the current government of the state, but to the enduring common life itself. Concerted disinformation campaigns seeking to foment discord or promote irredentist and secessionist aspirations run the risk of dissolving the common life entirely. Corrosive falsehoods promoted by such attacks may work to undermine the generations of shared experience which Walzer posits are integral to the manifestation and preservation of a common life. In more extreme cases, these efforts serve to subvert the very basis of common understanding necessary for cooperation between communities within the state. The net effect of these attacks is a common life that is, at best, inorganically altered by direct foreign interference or, at worst, irrevocably fractured as the diverse groups which comprise the state are left with no common ground upon which to build a cooperative collective identity. Accordingly, insofar as the state's inherent right to self-defense is

[237] Adela Suliman, "Sweden sets up Psychological Defense Agency to fight fake news, foreign interference," *The Washington Post*, January 6, 2022. Accessed January 16, 2022. https://www.washingtonpost.com/world/2022/01/06/sweden-fake-news-psychological-defence-agency/.
[238] Suliman, "Sweden sets up Psychological Defense Agency to fight fake news, foreign interference."

motivated by its right to unfettered political self-determination, cyberattacks posing an existential threat to the common life constitute just cause. As with Smith's Sovereignty View, considering these kinds of cyber operations as reasonable grounds for just cause is further bolstered by the remaining JWT criteria, ensuring that disinformation efforts which are trivial in scope and influence ultimately fail to motivate an outbreak of armed hostilities. Incorporating attacks which work through manipulation into the just cause framework marks such attacks as operations potentially severe enough to warrant a military response, offering target states some recourse in terms of formulating a substantial response and motivating aggressors to reconsider the utilization of these tactics.

4.6 Cyber *Casus Belli*

Insofar as the inherent right to self-defense stems from the need to protect the state's common life from foreign aggression, an evaluative account for just cause within cyberspace must necessarily be capable of accounting for the unique and ever-emerging threats cyber operations pose to the common life. While Walzer identifies that the crime of aggression arises from the violation of the rights of citizens, he ultimately limits potential acts of interstate aggression to boundary crossings and physical assaults. Although this distinction is sufficient for conventional uses of force, Whetham's subsequent account of cyber *chevauchées* illustrates how cyber operations without a physical footprint may nonetheless work towards accomplishing foreign policy objectives which had previously been reliant on the deployment of military forces, showing that cyber acts are readily capable of violating a state's common life. As interstate conflict moves into a new domain with hostile action between states now predominantly taking place within cyberspace, it becomes evident that we require a cyber-specific evaluative framework for just cause.

Noting the shortcomings of the Standard View's reliance on the presence of harms analogous to those of kinetic action, Smith instead posits a Sovereignty View of just cause. The Sovereignty View shifts the evaluative focus away from the physical consequences of cyber operations and towards the specific nature of *how* they seek to accomplish their objectives. Those cyberattacks which simply impose a cost meant to motivate a state towards a particular

outcome fail to fulfill the just cause criterion, while the more invasive attacks ensuring an outcome in spite of the organic deliberations of a target state satisfy the just cause criterion. Insofar as such attacks directly undermine a state's capacity to practice political self-deliberation, they serve as just cause. I further argue that a third type of cyberattack, that which seeks to achieve its objective through direct manipulation of a state's common life, likewise serves to satisfy the demands of just cause. Provided a cyber operation pursues a specific political objective by manipulating the common life to shape political deliberation in an inorganic fashion or to cripple a state's ability to politically self-determine due to a fractured common life, it fulfills the just cause criterion. While this loosens the demands of the just cause criterion and broadens the scope of cyberattacks which may reasonably satisfy it, this is tempered by the remaining *jus ad bellum* criteria which work to prevent the frivolous deployment of military force in response. Let's turn to those remaining criteria now.

Chapter 5

Right Intention and Public Declaration by Proper Authority

5.1 Tempering the Revised Sovereignty View

The move to forgo the Standard View in favour of adopting a revised Sovereignty View as the preferred evaluative approach for just cause in cyberwar removes arguably the biggest *jus ad bellum* hurdle for cyber operations. No longer does the absence of physical harm necessarily omit disanalogous cyberattacks from deliberations of just cause. Rather, this shift allows for the severity of such cyberattacks to instead be measured by the invasiveness of their approach and the degree of harm they are capable of inflicting against a state's common life. Under the revised Sovereignty View, a greater number of cyberoperations qualify as morally justified grounds for initiating a forceful response. Insofar as this approach broadens the spectrum of cyber operations realistically motivating the just cause criterion, further attention must necessarily be paid to the remaining *jus ad bellum* principles to temper the likelihood that any given case of just cause may escalate into full-blown interstate armed conflict. While most of the existing literature regarding *jus ad bellum* and cyberwar revolves around the just cause dilemma, notably less attention is given to the remaining principles and how they manifest within the cyber domain. This is ultimately problematic as, much like the just cause criterion itself, the transition into the digital domain introduces new considerations for these principles, some of which may undermine our existing preconceptions regarding their applicability. As such, in the interests of structuring a more complete *jus ad bellum* framework for cyberwar, it is necessary to consider how each of the remaining principles apply within cyberspace. Both this chapter and the one after seek to analyze the remaining *jus ad bellum* criteria to address their role within cyber conflict.

While the just cause criterion spearheads the JWT approach to *jus ad bellum*, it is joined by two further anti-consequentialist principles in the form of right intention and proper declaration. The principle of right intention demands that any mobilization of force be motivated by the pursuit of a benevolent cause, namely those specified by the just cause criterion.[239]

[239] Seth Lazar, "War", *The Stanford Encyclopedia of Philosophy* (Spring 2020 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/spr2020/entries/war/. Accessed Jan 25, 2022.

Broadly speaking, it is this principle which ensures that a state's motives for going to war are justified, rather than threadbare excuses meant to disguise the opportunistic advancement of foreign policy objectives. While the principle of right intention polices the motives of states, the principle of public declaration seeks to establish accountability within conflict. To this end, the principle demands that any forceful action taken by a state be predicated by a public declaration by the appropriate state authority.[240] This principle both ensures that the belligerents are appropriately identified and that the impetus behind their use of force is made apparent. Once the just cause condition has been deemed met, these two supplementary principles work in tandem to determine whether the ensuing forceful retaliations are motivated by the pursuit of morally justifiable objectives and whether the proper responsibility is taken by the appropriate parties. Although our experience with conventional conflict has resulted in certain emergent expectations with regards to these two principles, their translation into the digital domain should nonetheless be supplemented with further interpretation so as to account for the nuance of cyber-specific operations more fully.

5.2 War to What End?: Right Intention and *Jus ad Bellum*

Although the principle of just cause determines whether a state possesses a morally justified impetus for going to war, it alone is insufficient for establishing whether the specific nature of a state's response is likewise justified. Without further considerations, a fulfilled just cause criterion can lend moral credence to offensive campaigns with primary objectives that are aggressive and expansionistic. Supposing it was the victim of a series of cross-border raids launched by its neighbour, it would be uncontroversial to claim that State A has moral justification for deploying defensive force against State B to deter further aggression. It is more difficult to assert that State A enjoys the same moral justification should it use the raids as pretense for an expansionistic military campaign designed to annex large swathes of State B's territory. It is not an uncommon phenomenon for belligerent states to hide their true intentions behind appeals to just cause. The Nazi invasion of Poland in 1939 was itself predicated by a series of false flag operations, culminating in the Gleiwitz incident within which SS troops

[240] Brian Orend, *War and Political Theory*. (Cambridge, UK; Medford, MA: Polity, 2019), 88.

disguised in Polish military uniforms staged an attack on a German radio station near the border, intending to portray Germany as the victims of their neighbour's aggression, thus ascribing them *casus belli* for military action.[241]

While not a principle specified by the *jus ad bellum* of the laws of armed combat, the principle of right intention is added into the JWT framework to further constrain state decisions to respond to incurred aggression with war. In accordance with this principle, just cause alone is insufficient for morally justifying the deployment of a state's military force. Rather, the principle further demands that any such course of action be motivated by the right intent. This principle has its roots firmly in the work of Augustine who, seeking to reconcile the pragmatic necessity of war in the twilight of the Roman Empire with the general aversion to violence expressed in the New Testament, concluded that war was only justifiable in the interests of protection and self-defense against aggression: under Augustine's conception, a state goes to war not out of desire for conquest or glory, but rather out of love for its citizens and its desire to protect them from external threats.[242] This line of reasoning motivates Augustine's subsequent assertion that "it is an established fact that peace is the desired end of war".[243] A state concerned with the welfare of its citizens should strive to restore peace, rather than sow instability, as a prolonged state of conflict renders it more likely that harm may befall its people. Although going to war undoubtedly proves necessary in certain circumstances, it is not an undertaking to be actively pursued by a ruler or state. In cases within which war is unavoidable, the intent behind forceful action should be the restoration of a state of peace, rather than any opportunistic desire for territorial expansion or the subjugation of foreign peoples.

Walzer's own deliberations pertaining to right intention follow a similar trajectory to those of Augustine, albeit they are developed further to argue in favour of a specific conception of post-war peace. Despite continuing to posit that the restoration of peace ought to be the motivating intention behind a state's mobilization to war, Walzer takes exception to the notion that such a peace manifests as a return to the pre-war state of affairs; while this would admittedly

---

[241] M.R.D. Foot, "Conditions Making For Success and Failure of Denial and Deception: Democratic Regimes", In *Denial and Deception: The Twenty-First Century Challenge*, eds. Roy Godson & James J. Wirtz (New Brunswick, N.J: Transaction Publishers, 2002), 100.
[242] Orend, *War and Political Theory,* 90.
[243] Gary J. Bass, "Jus Post Bellum", *Philosophy & Public Affairs* 32, no. 4 (2004): 384–412. http://www.jstor.org/stable/3557994. 387.

restore some semblance of peace, it is unlikely to be a lasting peace insofar as such an outcome would mark a return to the exact set of circumstances that had led to the outbreak of hostilities in the first place.[244] For example, a populist head of State A may seek to channel domestic discontent amongst citizens against neighbouring State B, painting the latter as being singularly responsible for State A's weak economic standing. An act of aggression by State A may be militarily repelled by State B, however a subsequent peace agreement may prove precarious insofar as the same conflux of factors which led to war in the first place (such as inflammatory populist rhetoric, a sizeable military equal to that of State B, and poor domestic economic opportunities) remain unaddressed. This may lead to subsequent flare-ups of hostilities between these states, perhaps being made even more likely by further animosity having been generated during the conflict between them.

Accordingly, the correct intention of war is not merely the restoration of the state of affairs preceding the outbreak of conflict—the proverbial "*status quo ante bellum*"— but rather something Walzer terms "restoration plus." Orend states that "restoration plus" is, put simply, "a *more secure possession of our rights*, both individual and collective".[245] The objective of a just war must necessarily be the reaffirmation of individual and collective rights, both within the responding state *as well as* for the citizens of the aggressing state. It is insufficient for the purposes of restoration plus to simply rout an aggressor militarily; the resolution of armed conflict ought to entail further steps to reinforce the basis of individual and collective rights, as well as to address the factors which had led to their violation in the first place. Often, these efforts take the form of demilitarizing the aggressor state and imposing punishment on relevant parties so as to mitigate the threat they pose. These efforts may be supplemented with "political rehabilitation" initiatives to address potentially dangerous ideologies posing a high risk of future threats,[246] such as Nazism in the immediate aftermath of WWII. This is not to say that restoration plus demands that the responding party seize full control of the aggressor state. Rather, the culmination of armed conflict should be followed with active efforts towards restoring the aggressor state "as an independent political community, enjoying political sovereignty and

[244] Brian Orend, "Justice after War," *Ethics & International Affairs* 16, no. 1 (March 2002): 43–56, https://doi.org/10.1111/j.1747-7093.2002.tb00374.x. 45.
[245] Orend, "Justice after War", 45.
[246] Orend, "Justice after War", 47.

territorial integrity".[247] Through excising the root causes of a state's aggression and subsequently restoring the aggressor's political sovereignty, a responding state not only curbs current aggression, but works actively towards mitigating the likelihood of future aggression. The result is a stronger foundation for a lasting peace.

While restoration plus offers a clearly defined right intention for war, critics may express doubts regarding the inclusion of the principle of right intention within *jus ad bellum* on the grounds that intention is notoriously difficult to accurately evaluate. Although deliberations regarding whether an act fulfills the principle of just cause may be rife with disagreement within the international community, the potentially aggressive act is nonetheless empirically observable. Controversy may arise regarding thresholds and the act's severity, however material facts related to the act, such as its origin and methodology, are accessible to external observers. Even covert operations with minimal footprints can eventually be dragged into the light through thorough investigations. While the final evaluation of the severity may vary between observers, external parties nonetheless have access to a set of incontrovertible material facts from which they may form their conclusions regarding whether the just cause criterion is met.

Critics may assert that this degree of epistemic access is notably absent in the case of intention. While material facts pertaining to uses of force are externally accessible, the same does not hold true for the intentions of a belligerent state. Provided a state has incurred an observable act of aggression qualifying as just cause, it may simply assert that its subsequent mobilization for war has the sole intention of curbing further foreign aggression, with any wartime occurrence suggesting a contrary objective being dismissed as merely incidental. In response to an interstate act of aggression, State A may publicly declare that its intentions in going to war with State B align with the objectives outlined by Walzer's restoration plus, while internally seeking to use the moral justification offered by the just cause principle to instead launch an expansionistic campaign designed to seize strategic ports necessary for its long-term foreign policy objectives. The general epistemic opacity of intention could then be used to rebuff international criticism of State A's action. For example, State A may insist that the seizure of these key ports is a necessary step towards the (partial) demilitarization of State B required for the restoration of lasting regional peace.

[247] Orend, "Justice after War", 47.

There is no shortage of historical examples of states masking their true self-interested intentions with public declarations expressing more benevolent motivations. Many acts of aggression attributed to Russia since the dissolution of the USSR have been predicated with Russian assertions that its military deployments were conducted in pursuit of morally justified ends. The 2008 Russo-Georgian war was preceded by a lengthy Russian campaign of "passportization", the issuing of Russian passports to residents of former Soviet countries, particularly those residing in unrecognized breakaway states such as Abkhazia and South Ossetia, which Georgia recognized as part of its own territory.[248] As Georgia's government sought to grow its ties with NATO and hostilities between the Georgian military and separatists within Abkhazia and South Ossetia flared up, Russian troops invaded South Ossetia with the stated intent of "protect[ing] citizens of the Russian Federation from the Georgian offensive".[249] Despite Russia's appeal to humanitarian pretenses, its involvement in the Georgian conflict has been seen as an effort to maintain its influence within the South Caucasus.[250] A similar phenomenon took place during Russia's invasion into the separatist-held Donetsk and Luhansk regions of Ukraine in early 2022; predicated by publicly recognizing these two regions as independent, Russia insisted that its troops were being deployed to provide "peacekeeping functions" within the area, rather than constituting an invading force.[251] Further questions surround the United States' true motivations behind its invasion of Iraq in 2003, as its initial declared intention, that of pre-emptively removing the threat of Iraqi weapons of mass destruction (WMDs), failed to hold up to subsequent scrutiny as it became clear that the US had little to no evidence of such WMDs existing prior to launching its offensive.[252] These kinds of events lend credence to the argument that the principle of right intention seems minimally

[248] Emil Aslan Souleimanov, Eduard Abrahamyan, and Huseyn Aliyev, "Unrecognized States as a Means of Coercive Diplomacy? Assessing the Role of Abkhazia and South Ossetia in Russia's Foreign Policy in the South Caucasus," *Southeast European and Black Sea Studies* 18, no. 1 (January 2, 2018): 73–86, https://doi.org/10.1080/14683857.2017.1390830. 80-81.
[249] Souleimanov et al., "Unrecognized States as a Means of Coercive Diplomacy? Assessing the Role of Abkhazia and South Ossetia in Russia's Foreign Policy in the South Caucasus," 82-83.
[250] Souleimanov et al., "Unrecognized States as a Means of Coercive Diplomacy? Assessing the Role of Abkhazia and South Ossetia in Russia's Foreign Policy in the South Caucasus," 83.
[251] Anton Troianovski, "Moscow orders troops to Ukraine's separatist regions after Putin recognizes their independence," *The New York Times.* February 21, 2022. Accessed February 23, 2022. https://www.nytimes.com/live/2022/02/21/world/ukraine-russia-putin-biden/moscow-orders-troops-to-ukraines-breakaway-regions-for-peacekeeping-functions.
[252] Robert Jervis, "Reports, Politics, and Intelligence Failures: The Case of Iraq," *Journal of Strategic Studies* 29, no. 1 (February 2006): 3–52, https://doi.org/10.1080/01402390600566282. 37.

restrictive in *jus ad bellum* deliberations insofar as we lack direct access to the true intentions of belligerent states and states have little difficulty asserting a morally justified intent while covertly operating in pursuit of a completely different, in some instances perhaps even contrary, objective.

While it is undeniable that states have historically held their full intentions close to the chest, the principle of right intention is not contingent on taking a belligerent's stated intentions on faith alone. Rather, there are two ways within which a state's true intentions may be reasonably ascertained without immediate access to its internal deliberations. In the first, state intent may be evaluated rationally prior to the state undertaking action. Orend notes that "[i]ntentions are neither infinitely redescribable nor irreducibly private: they are connected to patterns of evidence, as well as constrained by norms of logical coherence".[253] Geopolitics does not take place in a vacuum. Contemporary foreign policy is influenced by a rich history of political, cultural, economic, and military interchanges between states. A state with a history of threadbare justifications for armed interventions which regularly culminate in long-term occupations reasonably engenders greater skepticism regarding its true intentions than a state with a history of mobilizing solely towards supporting multilateral peacekeeping initiatives. Similarly, greater scrutiny might be directed towards acts of interstate hostility between nations with a recent history of ethnic conflict, such as those which broke out between disparate cultural groups within the Balkans during the dissolution of Yugoslavia.[254]

In a contemporary example, the consistent post-Cold War efforts of the Russian Federation to project influence into former Soviet Bloc states motivates greater scrutiny of the true motives behind its increased military activity within Ukraine in early 2022. While Russia claims to be acting in the interests of self-defense and international security, other states openly voice their skepticism due to Russia's recent history of interference within the Crimean Peninsula, as well as its general rhetoric regarding Ukrainian sovereignty.[255] Accordingly,

---

[253] Brian Orend, *Michael Walzer on War and Justice*, (Montreal; Ithaca, [N.Y.]: McGill-Queen's University Press, 2000). 95.
[254] Victor Roudometof and Roland Robertson, *Nationalism, Globalization, and Orthodoxy: The Social Origins of Ethnic Conflict in the Balkans*, Contributions to the Study of World History, no. 89 (Westport, CT: Greenwood Press, 2001). 238.
[255] Anton Troianovski, Roger Cohen, & Katie Rogers, "Putin Warns the West and Ukraine but Keeps His Intentions a Mystery," *The New York Times*, February 7, 2022. Accessed February 9, 2022. https://www.nytimes.com/2022/02/07/world/europe/putin-macron-russia-france-ukraine.html.

history may be coupled with available contemporary evidence when gauging the veracity of a belligerent's stated intent. This history may either offer support for a state's claims should they align with its character, or it may promote greater scrutiny should the asserted intent represent a notable break from a pattern of past behaviour. In either case, external observers are capable of forming reasonable conclusions regarding whether a state's declared intent is genuine, or if there are stronger motivating factors underpinning its behaviour.

Secondly, a state's conduct once war is underway often betrays its true intentions. A state's actions within an armed conflict are readily observable by external parties. Everything from the tactics a state employs to the objectives it prioritizes throughout the hostilities offers insight into a belligerent's true intentions. Attacks launched against military installations are compatible with the end goal of removing an aggressor's ability to commit further acts of aggression, lending credence to a stated intent of self-defense. By contrast, a combatant which, having invaded a neighbouring state under humanitarian pretenses, prioritizes the seizure of historically contested territory or natural resources would rightly draw greater scrutiny of its true intentions for mobilizing for war. Beyond larger objectives, the specific tactics employed by combatants lend further insight into their underlying intentions. Orend suggests that certain tactics serve as immediate signs of states harbouring more sinister intentions; atrocities such as ethnic cleansing, massacres, and mass rape campaigns are all markers of conflict fueled by ethnic hatred, rather than incidental occurrences in the pursuit of a morally justified objective.[256] These are not acts conducted by states operating with morally justified intentions. Instead, these tactics actively undermine the objective outlined by Walzer's restoration plus by flagrantly disregarding and violating the inherent rights of individuals and collectives. Irrespective of what a state's claimed intentions for going to war may be, committing these kinds of grievous rights violations renders it less likely that a lasting peace with rights may feasibly be established.

Evidently, evaluating the principle of right intention through a purely *jus ad bellum* lens is overly myopic. Insofar as a state's behaviour within the state of war serves as a further barometer for its intentions, evaluative efforts should be extended to include the *jus in bello* stage of war. The same holds true for the *jus post bellum* phase of conflict as well; just as a state's behaviour within a state of war offers insight into its true intentions, so too do its post-war

[256] Orend, *Michael Walzer on War and Justice*, 95.

resolutions. Following the culmination of hostilities, the principle of right intention may further evaluate whether the belligerent's overall postwar strategy has led to, or at least made a reasonable effort towards, a more secure foundation for rights as prescribed by Walzer's restoration plus. State A invading State B to overthrow the latter's dictatorial government under humanitarian pretenses may rightfully face great scrutiny if, following the culmination of open conflict, it simply retreats within its territory and leaves State B war-torn and destitute. By contrast, postwar policies designed to help rebuild critical infrastructure, establish domestic security, and restore State B's political autonomy would facilitate a transition into a more secure peace in the aftermath of conflict, lending greater credence to State A's claimed humanitarian intention. The continued relevance of the principle of right intention throughout the three stages of conflict has motivated more holistic approaches towards gauging right intent within interstate hostilities.

Building on Kant's earlier work on JWT, Orend suggests that the principle of right intention should necessarily entail an explicit commitment by the state "both publicly and in advance, as a matter of right intention, to adhering to the other rules of war, contained in *jus in bello* and *jus post bellum*" rather than solely a preemptive assertion of its intent.[257] With regards to *jus in bello*, this approach supplements the traditional principle of right intention within *jus ad bellum* with article 6 of Kant's preliminary articles for perpetual peace, asserting that "a state must not use such treacherous methods as would destroy that confidence which is required for the future establishment of a lasting peace".[258] A state's behaviour in war directly impacts the likelihood of peace in the aftermath of hostilities. Systemic war crimes actively undermine restoration of a peace with rights by virtue of their flagrant disregard of the basic human rights of citizens within a combatant state. Similarly, a state reneging on ceasefires and other agreements between belligerents in pursuit of a tactical advantage serves to violate trust between combatants, rendering it less likely that an amicable agreement between the two is reachable. These kinds of tactics prove ruinous for efforts at ending war with minimal casualties. Instead, such methods

[257] Orend, *Michael Walzer on War and Justice*, 94.
[258] Brian Orend, "Kant's Just War Theory." *Journal of the History of Philosophy* 37, no. 2 (Apr 01, 1999): 323. http://search.proquest.com.proxy.lib.uwaterloo.ca/scholarly-journals/kants-just-war-theory/docview/1297333042/se-2?accountid=14906. 350.

render it more likely that a war will end only following a prolonged campaign of hostilities attempting to force an unconditional surrender.

Koeman notes that asymmetric conflict, such as the ongoing Global War on Terror (GWOT), poses another challenge along these lines. In instances of asymmetric warfare, enemy combatants are often poorly defined and difficult to identify by design, enabling them to efficiently conduct guerilla campaigns and punch above their weight, so to speak. This may lead to an occupying military force coming to treat the entire civilian populace as, first and foremost, potential enemy combatants; this approach may foster resentment amongst the occupied populace, contributing towards longer-term animosity and, in extreme cases, further outbreaks of violence.[259] Accordingly, a preceding commitment towards the right intention of establishing a lasting peace would necessarily dictate how states may conduct themselves during military action. It would necessarily preclude any violations of the existing LOAC to maintain a degree of trust required for potential peace agreements between combatants. Similarly, in the case of asymmetric warfare, it would necessarily shape how states deploy and build relationships with the civilian populace of an occupied state.[260]

This pre-commitment entailed by right intention should similarly be extended to include the commitment to post-war conduct conducive towards the establishment of peace. Any mobilization for war should be reinforced with post-war planning to restore the critical societal institutions necessary for a defeated combatant's political self-determination, as well as the vital infrastructure necessary for it to ensure its citizens' basic human rights. Beyond these immediate efforts, a pre-commitment towards restoring peace should place constraints on post-war settlements. States should not approach the post-war settlement process as an avenue by which they may legally exact revenge upon a defeated state.[261] While punishment of responsible parties may prove necessary following armed conflict, states should resist any inclination to unduly burden a surrendered state. The right intention to war, according to Walzer, is to bring about a more desirable state of affairs so as to avoid similar conflict in the future; indiscriminate

[259] Annalisa Koeman, "A Realistic and Effective Constraint on the Resort to Force? Pre-Commitment to *Jus in Bello* and *Jus Post Bellum* as Part of the Criterion of Right Intention", *Journal of Military Ethics* 6, no. 3 (September 2007): 198–220, https://doi.org/10.1080/15027570701585373. 204.
[260] Koeman, ""A Realistic and Effective Constraint on the Resort to Force? Pre-Commitment to *Jus in Bello* and *Jus Post Bellum* as Part of the Criterion of Right Intention", 204.
[261] Koeman, ""A Realistic and Effective Constraint on the Resort to Force? Pre-Commitment to *Jus in Bello* and *Jus Post Bellum* as Part of the Criterion of Right Intention", 204.

punishments against a state and its citizens, ranging from blanket economic sanctions to a stranglehold on their ability to politically self-determine, are likelier to sow further unrest. In severe cases, an unfair postwar settlement may even offer the punished state just cause for continuing to fight.[262] As a result, the right intention criterion should not be deemed fulfilled by a *jus ad bellum* declaration of a state's intent; rather, it should further entail a pre-emptive commitment to conduct conducive towards achieving restoration plus in both the *jus in bello* and *jus post bellum* stages of war. If a state is unwilling to commit to adhering to these prescribed rules of conduct, then, as Orend suggests, it "should never involve itself in such morally serious business as warfare".[263]

At this stage it is important to acknowledge that war is complex and that state intentions are rarely straightforward. Walzer himself asserts that a "pure good will" in global politics is often little more than an illusion.[264] Although states may enter conflict with the genuine intent of fostering lasting peace, it may not necessarily be, and indeed it rarely is, their only intention. The US-led coalition effort to expel the Iraqi military from Kuwait in 1991 may have been motivated with the intent of restoring regional peace, however Orend notes that this does not preclude further motives from feasibly being in play, ranging from a desire to secure Kuwaiti oil supplies to conducting a show of force meant to establish US military superiority during the dissolution of the USSR.[265] Similarly, NATO's 1995 intervention in Bosnia and Herzegovina was framed as a response to continued aggressive conduct by Serb forces, motivated in part by the massacre at Srebrenica earlier that year; however, critics suggested that the intervention was seen by NATO primarily as a means of projecting its influence into Europe, with the operation's humanitarian pretenses serving as merely justification for getting involved rather than the primary impetus.[266] While in both cases the publicly declared intent was bringing an end to aggression, it may not be the case that the asserted intent was the sole motivation for going to war.

The presence of ulterior motives alone is insufficient for determining that a state's move to war is unjustified. Rather, Walzer conceives of going to war as permissible provided states

[262] Orend, "Justice after War", 56.
[263] Brian Orend, *The Morality of War*, (Peterborough, Ont.: Broadview Press, 2006). 48.
[264] Orend, *Michael Walzer on War and Justice*, 94.
[265] Orend, *Michael Walzer on War and Justice*, 94.
[266] David N. Gibbs, *First Do No Harm: Humanitarian Intervention and the Destruction of Yugoslavia.* (Nashville: Vanderbilt University Press, 2009). 164-165.

have a genuine intent to end aggression and achieve restoration plus, irrespective of whether this is the sole motive. As Orend suggests, Walzer's pragmatic approach towards right intention renders it possible "to criticize some of the non-moral motives that states can have in going to war while still endorsing the moral motive".[267] This acknowledgement addresses the historically grey area of intent in war, as instances of states moving to war in pursuit of selfless objectives are comparatively few and far in between. Instead of holding states to a standard of intent only achieved with purely benevolent moves to war, Walzer asserts that it is sufficient for the morally justified motive to be present.[268] This weaker threshold for right intention serves a pragmatic purpose; restricting states from going to war if they possess any ulterior motives renders it less likely that states would ever engage in humanitarian interventions. Coalition efforts in Kuwait in 1991 and in Bosnia & Herzegovina in 1995 may arguably have been conducted in part due to ulterior motives, however these ulterior motives may have proved more motivating than a purely benevolent intent would have been. Wars are costly, both politically and financially. As such, wars conducted solely for the benefit of the citizens of a foreign state may not garner the same support as wars which simultaneously advance one's own foreign policy objectives in some capacity. Allowing for states to harbour further motives *in conjunction with* the morally justified objective of restoration plus renders it likelier that such humanitarian interventions may take place.

Although Walzer's conception of right intention leaves room for ulterior motives, it does not pave the way for states to operate with impunity. An intent to foster a secure, lasting peace necessarily precludes certain kinds of ulterior motives. Intentions to forcibly annex territory or seize the natural and economic resources of another state are generally inconducive to restoration plus. Likewise, a prolonged military occupation in the interests of controlling a state's political character is likely to eventually result in renewed outbreaks of hostilities. Just as certain kinds of tactics actively undermine the end goal of secure peace with rights, so too do certain motivations work towards continued instability. Despite Walzer's framework for right intention permitting ulterior motives, the necessary commitment to restoration plus places constraints on what kinds of ulterior motives may feasibly accompany right intention. Any ulterior motive cannot work

[267] Orend, *Michael Walzer on War and Justice*, 94.
[268] Orend, *Michael Walzer on War and Justice*, 94.

towards undermining the establishment of a more secure peace within which rights are protected against further aggression.

The principle of right intention tempers the moral justification of a state's decision to go to war once the just cause criterion is met. While a fulfilled just cause criterion is a necessity for a state to be justified in going to war, the principle of right intention further demands that the retaliating state's motives are themselves aligned with the morally justified goal of preventing further aggression and achieving a peace within which basic rights are better secured than in the period preceding conflict. Although at present the principle of right intention is largely evaluated as a part of the *jus ad bellum* framework, it remains pertinent throughout the *jus in bello* and *jus post bellum* stages of conflict insofar as a state's conduct, both during war and in its aftermath, is indicative of its true intent. The spirit of the principle of right intention necessitates a broadening of scope to include an additional constraint in the form of a pre-conflict commitment to operating in a manner which adheres to the further demands of *jus in bello* and *jus post bellum* in order to strengthen the likelihood of a peaceful resolution to conflict. Working in tandem with the just cause criterion, the principle of right intention seeks to establish not only that a state is justified in going to war, but also that the state's intentions are morally justified, preventing an initial just cause from being merely an excuse to subsequently engage in conflict in pursuit of solely self-interested objectives.

5.3 Actions Louder Than Words: Tactics Betraying Intent

The principle of right intention within the *jus ad bellum* stage of cyber conflict remains largely unchanged when compared to its conventional counterpart insofar as there remains a demand on states to only engage in hostilities to advance morally justified ends. Where the key distinctions for right intention in cyberwar arise is within the *jus in bello* and *just post bellum* phases of armed conflict, as the novel methodology of cyberwar tactics necessitates a slightly different evaluative approach to how states conduct cyberwar. In the case of conventional warfare, a belligerent's stated intent may be scrutinized with reference to the tactics it employs over the course of conducting a war and the resolutions it seeks to put in place following the culmination of hostilities. A state launching kinetic attacks against the military infrastructure of

an aggressor can be reasonably taken as conduct supporting the end objective of preventing further aggression and restoring a stronger foundation of peace. By contrast, the seizure of territory and the accompanying displacement of certain ethnic groups under humanitarian pretenses would be rightfully criticized; such an action is more likely to result in renewed hostilities further down the road than it is in lasting peace. Insofar as this interpretation of the principle of right intention necessarily entails restricting certain types of tactics within traditional warfare, so too should a cyber principle of right intention differentiate between tactics conducive towards establishing lasting peace and those that actively undermine the foundations upon which such a peace could be built. This section will seek to identifying cyber tactics which evidence a state's commitment to restoration plus during conflict and following it, as well as those which are incompatible with the principle of right intention.

Provided the appropriate right intention for war is generally acknowledged to be the prevention of further aggression and the restoration of a lasting peace, the unique nature of non-physical cyberattacks may render certain cyber tactics preferable alternatives to conventional military operations. Kinetic action is inherently destructive. Air strikes on key targets typically result in significant physical harm in the form of property damage and bodily injury, even prior to accounting for the possibility of collateral damage which has a demonstrable effect on increasing the likelihood of subsequent violence.[269] Ground invasions result in similar harms, as well as the destruction, whether inadvertent or intentional, of infrastructure integral to the preservation of basic human rights, such as power grids, hospitals, and water filtration plants, as well as government institutions necessary for building peace in the aftermath. Even if physical attacks are backed by right intention, the destructive character of conventional military action poses challenges for restoration plus. Rebuilding efforts are costly, time-consuming, and not always effective at restoring a war-torn state. Roughly half of states recovering from conflict fall victim to further unrest within the decade following, as the demands of extensive civil reconstruction lead to widespread civilian discontentment and a higher likelihood of rampant corruption.[270] Prolonged military occupations meant to maintain peace and support rebuilding

[269] Luke N. Condra and Jacob N. Shapiro, "Who Takes the Blame? The Strategic Effects of Collateral Damage," *American Journal of Political Science* 56, no. 1 (January 2012): 167–87, https://doi.org/10.1111/j.1540-5907.2011.00542.x. 184-185.
[270] Fredrik Galtung and Martin Tisné, "A New Approach to Postwar Reconstruction," *Journal of Democracy* 20, no. 4 (2009): 93–107, https://doi.org/10.1353/jod.0.0132. 93.

operations may similarly foment resentment amongst the local populace and potentially trigger further unrest. Even in the best of cases, conventional action raises numerous post-war challenges even if the state is genuinely committed to restoration plus.

In this regard, a belligerent prioritizing cyber operations over kinetic alternatives may serve as evidence of its commitment towards restoration plus. As discussed within the first chapter, cyber operations are no longer niche activities deployed solely as elements of espionage campaigns. Rather, the continuing evolution of cyberweapons and tactics has resulted in cyber operations emerging as realistic alternatives to conventional action in many contexts. Stuxnet showed that cyberweapons are capable of remotely destroying state infrastructure, an objective which would have previously fallen under the sole purview of kinetic strikes. Similarly, cyber operations can diminish the military capabilities of belligerent states without resorting to inflicting wide scale physical destruction and loss of life. Cyber operators may target any number of military computer systems to undermine a state's ability to wage war, whether by hamstringing their logistics by rerouting key supplies or troop movements, overriding maritime and aerospace navigation systems, or even hijacking lethal autonomous weapons such as drones.[271] These efforts can collectively serve to reduce both a belligerent's ability and will to continue fighting, potentially motivating it towards negotiating a peaceful settlement to war.

Prioritizing cyber operations over kinetic alternatives wherever possible serves to leave a stronger foundation upon which to build peace due to two primary factors. Firstly, a predominantly cyber war reduces the necessity of a large physical footprint within a foreign state's territory. Occupations are rarely, if ever, straightforward. Occupying military forces must walk a fine line between building a relationship with civilians to foster cooperation and build positive sentiment, while simultaneously guarding against embedded combatants.[272] Mounting casualties amongst occupying forces may increase this tension, resulting in actions that may further aggravate the relationship between occupying personnel and local civilians. In extreme cases, this tension may result in atrocities. Perhaps the most infamous example may be drawn in the form of the My Lai Massacre in 1968. Prior to the massacre, the troops of C Company had

[271] Bruce Schneier and Tarah Wheeler, "Hacked drones and busted logistics are the cyber future of warfare," *The Brookings Institution*. June 4, 2021. Accessed February 14, 2022. https://www.brookings.edu/techstream/hacked-drones-and-busted-logistics-are-the-cyber-future-of-warfare/.
[272] Condra & Shapiro, "Who Takes the Blame? The Strategic Effects of Collateral Damage," 169.

engaged in a series of escalating acts of revenge against Vietnamese civilians motivated at least in part by growing casualties among the company during the Vietnam War.[273] These atrocities would eventually culminate in the American troops killing at least 347 Vietnamese civilians in the village of My Lai and subsequently reporting the incident as a successful operation against Viet Cong forces.[274] Such outbreaks of violence and heavy-handed responses by occupying forces serve to complicate both wartime- and postwar reconstruction efforts, rendering an already challenging process even more difficult. While it is likely that some physical presence is required when conducting a war between states, a heavier focus on cyber strategy may allow for a state to minimize its physical footprint in wartime hot zones.

Secondly, a predominantly cyber approach to conflict may mitigate the need for extensive physical postwar reconstruction of critical infrastructure. Kinetic strikes and conventional military operations run the risk of wide-spread destruction, causing lasting damages to key infrastructure even if such infrastructure is not itself a target. Once hostilities have ceased, there remains the daunting task of physically rebuilding damaged infrastructure. Postwar reconstruction projects face numerous complicating factors ranging from poor postwar planning by the victorious state to rampant corruption as opportunistic actors look to capitalize on the resulting turmoil within the rebuilding state.[275] In the case of Iraq, a significant number of Iraqi citizens still had limited access to electricity and drinkable water nearly a decade after the US' invasion in 2003.[276] In some conflicts, cyber weapons offer states an alternative tool towards achieving their objectives without the corresponding risk of extensive physical damages. For example, a state may seek to disrupt a belligerent's weapons manufacturing capabilities as part of its warfighting strategy. While an air strike may prove an effective approach, it may run the risk of causing collateral damage to a nearby power grid or hospital. By contrast, a cyberweapon may prove capable of shutting down the automated components of the manufacturing facility, functionally achieving the same objective, without the same risk of collateral damage. In the latter case, the more discriminate approach of the cyberattack results in less physical destruction and necessary rebuilding in the postwar period. Even cyberattacks against critical infrastructure

[273] Matthew Lippman, "War Crimes: The My Lai Massacre and the Vietnam War," San Diego Justice Journal 1, no. 2 (Summer 1993): 295-364. https://heinonline.org/HOL/P?h=hein.journals/tjeflr15&i=301. 300-301.
[274] Lippman, "War Crimes: The My Lai Massacre and the Vietnam War," 309-310.
[275] R. Jeffrey Smith, "The Failed Reconstruction of Iraq," *The Atlantic*, March 15, 2013. Accessed February 14, 2022. https://www.theatlantic.com/international/archive/2013/03/the-failed-reconstruction-of-iraq/274041/.
[276] Smith, "The Failed Reconstruction of Iraq."

deemed necessary for the war effort may be designed to be easily reversible by the attacker once hostilities have ceased, further reducing the need for lengthy reconstruction projects.

Insofar as both cyber- and conventional operations enjoy similar efficacy in achieving certain objectives, a belligerent's decision to deploy primarily cyber operations wherever possible may serve as evidence of it possessing the right intention within this context. By minimizing the deployment of kinetic military force, or forgoing it altogether, a state operates in a manner more conducive to the restoration of meaningful postwar peace due to such wartime operations being less physically harmful. By pursuing wartime objectives using non-destructive tactics, the subsequent task of rebuilding a state is rendered more manageable insofar as key infrastructure, ranging from utilities to roadways, will not have incurred the same degree of damages regularly entailed by conventional conflict. Furthermore, the deployment of transient cyberattacks may allow states to effectively achieve certain military objectives while still being easily reversible following the conflict resolution stage. The ability of cyberwar to achieve military objectives while causing less corresponding permanent harms allows for the state to emerge from the conflict with a greater degree of domestic security, rendering further civil unrest and subsequent outbreaks of violence in the postwar period less likely than if the state required extensive reconstruction efforts and a prolonged foreign occupation to assist in maintaining its internal stability. Provided that some military objectives may be achievable using solely cyber operations, the deployment of cyber measures over kinetic alternatives supports a state's claims of right intention, as such efforts may mitigate long-term, physically destructive damages and promote a more secure foundation for lasting peace.

This is not to say that a state's prioritization of cyber strategy over conventional alternatives is alone sufficient for determining that it is acting with the right intention. Just as certain kinds of cyber tactics suggest a correct motivating intention, so too do others betray a state's less benevolent motives. A state may seek to undermine a state's warfighting capabilities by deploying concerted disinformation campaigns designed to amplify domestic rifts and introduce political decision paralysis within an aggressor state. It may paint the current ruling party as being run by radical nationalists or controlled by foreign interests, potentially turning the citizens against the government. Such operations may result in the target state needing to fight a war on two fronts: an international one with a foreign belligerent, as well as a domestic one as it seeks to quell internal unrest amongst its citizens. In more extreme cases, these efforts may seek

to upset a state's internal stability by exacerbating underlying tensions between citizens of varying classes, ethnicities, political affiliations, and religious membership. History contains numerous atrocities fueled by inflamed divisions along these lines, including recent examples in the form of Serbian aggression against non-Serb civilians in Eastern Bosnia in the early 1990s,[277] and the 1994 Rwandan Genocide within which at least half a million Tutsi's were killed over the span of a month.[278]

It is undeniable that these kinds of divisive operations may be effective at reducing an aggressor's collective will to fight and accelerate an end towards interstate conflict. However, these operations may serve to shift the locus of conflict into a state's borders as the social fabric of shared society unravels. In more benign cases, these efforts may cause generalized animosity between groups, straining domestic political self-determination but not rendering it impossible. In more extreme cases, persistent efforts to undermine state unity and stratify a state's citizens can result in the outbreak of civil war with groups vying for control of the state government or breaking away in the interests of forming autonomous states. In either case, the conclusion is not a more secure foundation for peace. Rather, the rights of citizens face further threats as domestic tensions flare and the common life serving as the crux of the state falls apart. Accordingly, while these kinds of cyber operations may diminish a state's ability to act aggressively, they also serve to open a figurative Pandora's box of domestic volatility which renders long-term project of a lasting peace with rights more challenging, if not untenable.

Insofar as the principle of right intention entails a commitment to constraints on conduct within the *jus in bello* and *jus post bellum* stages of conventional war, it likewise places a commitment to constraints on permissible cyber conduct. Should we hold Walzer's restoration plus as the morally justified right intention for war, a state's commitment towards substituting cyber operations for conventional action *where it is reasonable to do so* expresses right intention. By forgoing physically destructive conventional attacks where effective cyber alternatives exist, a state expresses its commitment to bringing about a resolution to conflict which mitigates collateral damage and post-war unrest during the reconstruction period. A stronger reliance on

[277] Lara J. Nettelfield and Sarah E. Wagner, *Srebrenica in the Aftermath of Genocide* (Cambridge; New York: Cambridge University Press, 2014). 8-10.
[278] René Lemarchand, "The 1994 Rwanda Genocide," in *Century of Genocide: Critical Essays and Eyewitness Accounts*, eds. 3rd ed. Samuel Totten and William S. Parsons, (New York: Routledge, 2009). 420.

temporary or reversible cyber measures allows for the quicker restoration of basic rights following the cessation of hostilities. A stronger reliance on non-physically harmful cyber strategies may similarly alleviate some concerns that a state is conducting war primarily as a means of exacting a lasting revenge on an aggressive state, or to seize swathes of territory under pretenses of self-defense. The discriminate, non-physical, and temporary nature of certain cyber strategies uniquely positions them as tools towards stopping aggression, while reducing the threat war poses to the basic human rights of citizens within warring states.

By contrast, a cyber principle of right intention necessarily restricts cyber operations that actively work against the project of a more secure peace with rights. Purposefully inflammatory disinformation campaigns may serve to stop interstate aggression by reducing an aggressor's ability to wage war, however they do so at the cost of fomenting internal strife and unrest within an aggressor's borders. Such tactics are impermissible under the cyber principle of right intention as they actively work against the morally justified end objective of restoration plus. While interstate hostilities may end, the basic rights of the targeted state's citizens are placed in jeopardy as internal rifts are artificially amplified and conflict is turned inward, setting the stage for political turmoil, civil war, or, in the worst cases, genocide. The long-term effects of these operations render it less likely that we may achieve restoration plus in the aftermath of interstate conflict; rather, they instead ensure that the postwar period within the affected region will be rife with instability. As a result, the principle of right intention in the cyber context entails the prioritization of transient, less destructive cyber operations over physically harmful kinetic attacks wherever both options enjoy similar efficacy. Cyber right intention further prohibits the employment of destabilizing cyberattacks striving to create or amplify internal strife as a means of either reducing a belligerent's ability to wage war, or to advance one's own foreign policy objectives. We may now turn to the remaining anti-consequentialist principle of the *jus ad bellum* framework.

5.4 Public Declaration by Proper Authority

The anti-consequentialist principles of the *jus ad bellum* of the JWT tradition are rounded out by the principle of public declaration of war by a proper authority. While initially seeming

somewhat trivial, the principle of public declaration serves to fill in key gaps within the *jus ad bellum* framework not fully encompassed by the remaining principles. The principle of public declaration demands that war be "declared out in the open, officially and honestly, by the individual or government department/ministry with the authority for doing so".[279] Born of a fear that ambitious Roman generals may privately profit through their usage of the public Roman military, this principle seeks to introduce checks and balances for the declaration stage of war.[280] The inclusion of the principle ultimately ensures that any interstate conflict must be declared by a legitimate authority representative of the state, rather than merely a rogue element or individual within it.

The principle of public declaration of authority serves three primary purposes. Firstly, requiring public declaration issued by a proper authority works to ensure that the decision to go to war in some part reflects the political self-deliberations of a state's citizens. While the specific nature of the division of war power differs from state to state, in most democratic states the authority to declare war lies firmly within the jurisdiction of its legislature.[281] In democratic societies, the composition of the legislature is itself determined by the citizens. As such, potential decisions regarding mobilizations for war are ultimately deliberated on by a collective of representatives meant to act in accordance with the will of their constituents. Insofar as war is deleterious to the human rights of a state's citizens, it is imperative that the citizens "meaningfully consent to the launching of a war on their behalf".[282] Accordingly, a state which goes to war against the will of its citizens ignores their right to political self-determination.

The perceived appropriateness of a declaration of war is often reflected in how it is received by the constituents of the state making the declaration. The US' declaration of war against the Axis powers within WWII was greeted with strong domestic support, while its subsequent involvement in the Korean War less than a decade later failed to garner the same sort of widespread approval amongst US citizens; while a strong sense of the "righteousness of [their] cause" motivated the willingness of citizens to be mobilized for war in WWII, these sentiments

[279] Orend, *War and Political Theory*, 88.
[280] Orend, *War and Political Theory*, 88.
[281] Orend, *War and Political Theory*, 88.
[282] Orend, *The Morality of War*, 50.

were less present in the war immediately following.[283] The Vietnam War would prove even more controversial domestically, as evidenced by widespread domestic demonstrations and political unrest during the US' campaign in Vietnam. The difference in public sentiment towards these conflicts can be in part linked with the manner within which the wars were declared by the US. The 1941 decision to go to war with Japan was ultimately enacted by US Congress after the Pearl Harbour attacks dispelled the US' policy of isolationism and galvanized public sentiment against the Axis powers.[284] By contrast, the subsequent decisions for US involvement in the Korean War and the Vietnam War were both initially made by the executive branches of the government, rather than by Congress.[285] The principle of public declaration by proper authority seeks to mitigate the possibility that war is unilaterally initiated by a state *against the interests of its constituents*. To this end, a public declaration of war allows one's own citizens to voice whether any given war is in their best interests, whereas wars fought in secret may seek to subvert the political self-determination of a state's citizens entirely, leaving them little to no recourse.

This ties into the second purpose of the principle of public declaration; namely, to establish accountability, both domestically and abroad. By demanding that any state's decision to go to war must be publicly declared, this principle seeks to dispel any uncertainty regarding where states stand in relation to one another, as well as ensuring that any armed attacks conducted between states are appropriately attributed. Concerns regarding attribution have become particularly relevant as conventional warfighting tactics shift away from widescale pitched battles between amassed military forces, and instead move towards smaller scale engagement conducted primarily with precision deployments by special forces and remote warfare tactics such as drone strikes. As the average military operation grows more covert, the principle of public declaration takes on added importance due to the risks posed by unattributed attacks. As mentioned prior, covert action undertaken without subsequent public declaration may seek to specifically avoid the public scrutiny and deliberations of a state's citizens. This reduces their political self-determination insofar as the citizens' input regarding whether such operations

[283] Edward A. Suchman, Rose K. Goldsen, and Robin M. Williams, "Attitudes Toward the Korean War." *The Public Opinion Quarterly* 17, no. 2 (1953): 171–84. http://www.jstor.org/stable/2746273. 171.
[284] Inderjeet Parmar, "Catalysing Events, Think Tanks and American Foreign Policy Shifts: A Comparative Analysis of the Impacts of Pearl Harbor 1941 and 11 September 2001," *Government and Opposition* 40, no. 1 (2005): 1–25, https://doi.org/10.1111/j.1477-7053.2005.00141.x. 14-15.
[285] Orend, *War and Political Theory*, 89.

reflect their values is actively avoided by their state. While citizens may readily consent to military action conducted in the interests of immediate self-defense or humanitarian intervention on behalf of a beleaguered nation, they may prove much less willing to endorse manipulative intervention action meant to stymy the political self-deliberation of other states.

Beyond domestic accountability, the principle of public declaration works towards establishing international accountability. Admittedly, undeclared operations may present enticing propositions for states by offering avenues towards their foreign policy objectives while mitigating the risk of incurring heavy-handed punitive responses by international governing organizations. Notably, such operations need not necessarily be entirely covert. In some cases, the lack of a public declaration may do little more than offer states "implausible deniability", rendering their involvement an open secret that allows them to convey a political message, while still falling short of providing concrete grounds for full-fledged warfare.[286] As evidenced by Russia's 2014 hybrid warfare efforts in Crimea and Donbas, less covert action proves effective at "generat[ing] a situation where it is unclear whether a state of war exists—and if it does, who is a combatant and who is not".[287] Despite widespread reasonable belief of Russia's involvement with, and direction of, regional separatist movements, there remains enough ambiguity within the situation for Russia to avoid severe reprisals (at least up until 2022).[288] The obfuscatory effect of these operations not only renders conflict settlements more difficult due to a lack of transparency as to who the relevant stakeholders are, but it also makes the task of enforcing accountability amongst combatants more daunting. As evidenced by the lengthy investigation into the parties responsible for shooting down Malaysia Airlines Flight 17 within Ukrainian airspace in 2014, the existence of various state and nonstate actors within an uncertain conflict zone leads to a fraught environment within which responsibility for criminal acts is harder to discern than it would be within a formal state of war between two recognizable combatants.[289] These complications may be dispelled with public declaration of involvement by the relevant states.

---

[286] Rory Cormac and Richard J. Aldrich, "Grey Is the New Black: Covert Action and Implausible Deniability," *International Affairs* 94, no. 3 (May 1, 2018): 477–494, https://doi.org/10.1093/ia/iiy067. 488.
[287] Cormac and Aldrich, "Grey Is the New Black: Covert Action and Implausible Deniability," 490.
[288] Cormac and Aldrich, "Grey Is the New Black: Covert Action and Implausible Deniability," 490.
[289] Somini Sengupta and Andrew E. Kramer, "Dutch Inquiry Links Russia to 298 Deaths in Explosion of Jetliner Over Ukraine," *The New York Times*. September 28, 2016. Accessed February 16, 2022. https://www.nytimes.com/2016/09/29/world/asia/malaysia-air-flight-mh17-russia-ukraine-missile.html.

The third benefit of the principle of public declaration ties back into the principle of right intention. A state cannot merely assert that it is going to war without providing its motives. Public declarations of war require a state to make a public case for their military action by expressing their reasons for going to war and their general objective. This transparency in pre-conflict declarations is important insofar as it opens a state's decision to external scrutiny. State declarations of intention are sometimes presented in a "scatter-shot" manner, within which a state justifies their decision with appeals to a myriad of potential explanations in hopes of finding one that is deemed sufficiently acceptable, whether domestically or in the eyes of the international community.[290] The Unites States' offensive in Iraq in 2003 was one such case; prior to launching its offensive, The US sought to justify its engagement with various stated intentions from dismantling Iraq's suspected WMD capabilities, to humanitarian intervention on behalf of the Iraqi people suffering under a tyrannical government.[291] While more justificatory reasons may theoretically offer further support for a state's decision to go to war, Orend cautions that the scattershot method of declaring intention runs the risk of masking the true strength of the pro-war argument and contributing to the illusion that a state's basis for going to war is stronger than it may be in actuality.[292]

The principle of public declaration offers an avenue by which states may be forced to, as Orend suggests, "call their shot".[293] Orend proposes that it is insufficient for states to simply declare their decision to go to war, suggesting that an additional demand ought to be placed on states to also provide their "*main reason* for resorting to war".[294] This further constraint offers two primary benefits. Firstly, necessitating that a state be forthright in stating its primary reason for war offers a specific cause against which to measure a state's conduct in war. A scattershot justification for war may see a state launch a myriad of operations towards seemingly disparate objectives, rendering it difficult to ascertain whether the state is operating in a justified manner or if it is going beyond what is reasonable given its intentions. This leads into the second benefit of focus. Forcing a state to be transparent regarding its primary motivation for war encourages it to be more focused in its warfighting strategy. Having a key strategy in mind is conducive to

[290] Orend, *The Morality of War*, 49.
[291] Orend, *The Morality of War*, 49.
[292] Orend, *The Morality of War*, 49.
[293] Orend, *The Morality of War*, 49.
[294] Orend, *The Morality of War*, 49.

formulating a narrower approach to war that is more likely to achieve its objective than an unfocused effort towards multiple disparate objectives.[295] This may have the additional effect of minimizing unnecessary collateral damage over the course of a conflict.

Critics may suggest that the principle of public declaration is untenable insofar as an incurred act of aggression may require a timely response. The relatively slow pace of domestic political deliberation may mitigate the likelihood of a state being able to respond within a reasonable timeframe, rendering their eventual course of action less effective. Similarly, a declaration of a state's mobilization for war offers advanced warning for the aggressor state, allowing it further time to prepare for retaliatory measures, thereby reducing their potential impact. In either case, the principle of proper declaration seems to work to an aggressor's advantage. However, Orend suggests that this fear is overblown. The pragmatic concerns of war often lead to the division of war power across multiple levels of government. Accordingly, the executive branch may enjoy the authority to immediately respond to acts of aggression which demand swift responses. It is acceptable for in-depth democratic deliberation, as well as a public declaration of war, to take place after the fact, particularly if the conflict subsequently demands greater state involvement beyond the immediate response.[296] As a result, it is entirely possible for states to have mechanisms in place to facilitate rapid responses to aggression without precluding the principle of proper declaration.

Evidently, the principle of public declaration by a proper authority stands as more than a mere formality within *jus ad bellum*. It instead serves as a pragmatic addition to the preceding anti-consequentialist principles of just cause and right intention, demanding that any state's decision to go to war be followed by a public declaration rendering the state's justification and objectives readily apparent. In so doing, the principle works towards ensuring that a state's reasons for war align with the political interests of its state's citizens, as well as establishing wartime accountability both domestically and internationally. Furthermore, the principle seeks to ensure that states mobilizing for war are focused in their approach, rather than rushing into the heavy business of war with poorly conceived, and/or deceptively many, intentions.[297] With a

[295] Orend, *The Morality of War*, 49.
[296] Orend, *Michael Walzer on War and Justice*, 96.
[297] Orend, *The Morality of War*, 49-50.

foundation for the principle of public declaration established, we can now examine how the unique nature of cyber operations adds an additional dimension to this principle.

5.5 Dispelling the Attribution Problem

By and large, the principle of public declaration serves to achieve the same goals in the cyber domain as it does within the traditional domains of conflict. Where it takes on additional importance is at the level of attack attribution. Conventional military operations, whether covert or overt, are often readily attributable to the attacker within a reasonable time frame. Airborne attacks conducted by aircraft or drones, as well as shipborne cruise missiles, are trackable via the early warning systems of air defense grids, rendering it possible for a target to discern the attack's origin in a swift manner. Similarly, clandestine deployments of special forces troops may have a small footprint, however they are typically underpinned by an expansive network of supporting infrastructure allowing the sponsoring entity to project force abroad, such as forward operating bases and airfields or aircraft carriers; while attribution beyond a shred of doubt in such cases may be difficult by design, the targeted state can, with good reason, rightfully suspect the origin of the operation. As a result, conventional attacks across the spectrum of secrecy are often attributable within a reasonable timeframe.

By contrast, the specific axis along which cyberattacks work renders the task of attribution a trickier prospect. Cyberattacks often have no physical footprint. Barring cyberattacks which may require physical access to secured computer systems, cyber operations fail to constitute any sort of literal intrusion into a target's physical territory by virtue of being launched remotely. The methodology of certain cyberattacks further complicates matters. Botnet attacks are comprised of potentially thousands of computers previously compromised by malware and linked to a network which may be weaponized and deployed as part of a cyber operation; in such attacks, each of these "zombies" functionally serves as an individual attacker, masking the true origin of the operation.[298] The comparatively minimal infrastructure demands of cyber operations offers another layer of secrecy, as effective cyberattacks can be conducted

[298] W. T. Strayer, R. Walsh, C. Livadas, and D. Lapsley, "Detecting Botnets with Tight Command and Control," *Proceedings. 2006 31st IEEE Conference on Local Computer Networks*, 2006, pp. 195-202, doi:10.1109/LCN.2006.322100. 195-196.

without specialized infrastructure. Problematically, cyberattacks have hitherto often worked to hinder or shut down the network infrastructure of their targets; these kinds of attacks are readily capable of hitting exactly the sort of computer systems required to trace the origin of the cyberattack in a timely manner. This may render it impossible for a state to accurately discern an attacker's identity until long after the attack has subsided, typically on the attacker's terms. These concerns are further compounded by the continued applicability of traditional measures designed to shroud an attack's origin, such as the employment of proxies. As a result, the project of attribution grows more difficult with the transition into cyber space.

This is not to say that the majority of cyberattacks go completely unattributed. Specialized forensic teams can examine the methodology of an attack and discern with a reasonable degree of accuracy the party responsible. Problematically, these efforts are often time-consuming processes that may take weeks or months to come to their conclusions. In the case of Stuxnet, the attack's origin was only uncovered by a global team of investigators drawn from multiple cyber security companies after seven months of attempting to reverse engineer the worm.[299] Even so, it was estimated that Stuxnet had been operational for at least two years prior to the worm first being discovered, raising further attribution concerns regarding cyberattacks which are only detectable well after their initiation and only "when the evidence is deteriorating".[300] While the attribution problem is not absolute, the general complexity of discerning cyber attributability contributes to an unignorable disparity between conventional- and cyberattack attribution timeframes.

This attribution problem poses risks for states seeking to respond to incurred acts of cyber aggression. Aggression often necessitates swift responses by the victimized party, as targeted states may feel intense pressure domestically to defend their citizens and their own sovereignty against a hostile foreign entity. In the case of conventional attacks, a state may immediately respond militarily against the infrastructure used by the aggressor to conduct their aggressive operation. Likewise, they may enact swift sanctions against the responsible party while seeking further diplomatic resolutions. However, both such options are contingent upon a state's ability to efficiently identify the responsible party. In the case of cyber operations, the responding state

[299] Marcus Schulzke, "The Politics of Attributing Blame for Cyberattacks and the Costs of Uncertainty," *Perspectives on Politics* 16, no. 4 (December 2018): 954–68, https://doi.org/10.1017/S153759271800110X. 956.
[300] Schulzke, The Politics of Attributing Blame for Cyberattacks and the Costs of Uncertainty," 956.

is likely to be operating with comparatively poor certainty as to an aggressor's identity. Despite this, there is an understandable "temptation to blame the most obvious rival" in the immediate aftermath of any incurred act of aggression as tensions are high and the targeted state scrambles to defend itself.[301] However, just as we have access to a wealth of geopolitical history which we may use to evaluate the likelihood of a state's responsibility for any attack, so too can potential aggressors seize the opportunity provided by historical tensions to advance their own foreign policy objectives, reasonably believing another state may be held accountable for its covert operations.

In the cases of milder cyberattacks, the lack of definitive attribution in the immediate aftermath may lead to investigative efforts adopting a skewed evaluative perspective which largely seeks to reverse engineer a reason why the immediately suspected aggressor is the responsible party, rather than leaving open the possibility of a third-party attacker; Schulzke warns that a predisposition to immediately ascribe blame without definitive supporting evidence "helps to entrench initial frames and promotes confirmation bias, with new information being melded to the existing frame rather than being used to introduce alternative explanations". This may serve to inflame tensions needlessly between states if the traditional rival is not itself the aggressor in this case. In more extreme cases, a third-party aggressor may launch a destructive cyberattack amid a period of elevated tensions between two historically opposed states. Presuming such a cyberattack cannot be accurately attributed to the correct entity within a reasonable timeframe, the domestic demands placed on a victimized state to respond to foreign aggression may result in the wrongful attribution of the attack to its neighbour, potentially triggering full-scale conflict between states. While similar phenomena may occur within the realm of conventional warfare, the unique attributional challenges posed by cyber operations renders this a larger concern within the digital domain.

The principle of public declaration seeks to redress some of these concerns by promoting greater transparency regarding state cyber operations. Notably, a public declaration need not be necessary for every type of cyber operation. For example, there would be understandable hesitance by states to publicly declare their cyber espionage efforts insofar as declaring them would serve to dramatically undermine their effectiveness; as such, the likelihood of any state

[301] Schulzke, The Politics of Attributing Blame for Cyberattacks and the Costs of Uncertainty," 959.

admitting the scope of their cyber espionage campaigns is slim. However, I propose that any cyber operation which may reasonably be construed as fulfilling the just cause criterion must be accompanied by a public declaration by the proper authority. Insofar as certain classes of cyber operations run the risk of providing grounds for going to war, the states deploying them must necessarily declare their responsibility, as well as the intentions which had served to motivate the operation. This would ensure that the most serious cyber operations, those which may reasonably be responded to with force, are 1) attributable to the responsible aggressor, and 2) accessible to the citizens of the state so that they may determine whether such operations reflect their political interests.

Firstly, the principle of public declaration serves a key role in offsetting the higher probability of mistaken retaliation which plagues cyberwar to a stronger degree than conventional conflict. Just as with conventional operations, a cyber operation need not necessarily be preceded by a public declaration. It may be the case that declaring a retaliatory cyber operation against an aggressive state may serve to shift the target state's defensive priorities away from conventional military strength to greater cyber preparedness, reducing the effectiveness of an incoming cyberattack; as such, it is unlikely that states would commit to declaring imminent operations. However, the principle nonetheless demands that a state provide a public declaration *within a reasonable timeframe* following an attack. It is impermissible for states to launch comprehensive cyber uses of force without subsequently claiming responsibility for the attack. By claiming responsibility, a state renders itself accountable and makes it readily apparent that there exists a state of war between itself and its target. This works to reduce the likelihood of collateral damage arising due to mistaken attribution which may drag a third, otherwise uninvolved, state into conflict. By forgoing a declaration in hopes that the targeted state may either be crippled by indecisiveness or potentially take further retaliatory measures against an uninvolved party based on mistaken attribution, the initial aggressor breaks the provisions laid out within *jus ad bellum*. While the initial attack may not be traced back to them in the immediate aftermath, an eventual investigation and attribution may open such a state up to severe international sanctions or other punitive measures.

Secondly, public declarations of cyber operations ensure that a state's cyber campaigns fall under appropriate domestic scrutiny. Conventional operations are predominantly accessible to the public, as troop deployments and sorties can be tracked by a variety of watchdog

organizations, and emphatically with all of today's online- and social media tools. As such, citizens can gauge whether their state's military operations align with their self-determined political values. By contrast, cyber operations are typically conducted well out of the public eye. Should speculation emerge as to a state's responsibility, it is often the case that the cyber operation is shrouded with layers of plausible deniability. Even in cases within which a state's responsibility is certain, the idiosyncratic methodologies of cyberattacks render it likely that citizens are unaware as to the true scope or effect of their government's cyber operations. As such, the inherently covert nature of cyber operations may prove "a very seductive temptation for those with the war power".[302] This proves problematic insofar as incurred cyberattacks may reasonably constitute grounds for going to war, potentially resulting in severe ramifications for a state's citizens who may be wholly unaware of their government's aggressive posturing within the digital domain. Accordingly, the principle of public declaration seeks to assuage these concerns by ensuring that governments cannot wage secret wars across cyberspace against the political interests of their constituents. Rather, by forcing a state's cyber affairs into the light, the principle of public declaration allows for forceful cyber operations to be added into the fold of a state's general military capabilities and evaluated accordingly: as war measures akin to conventional attacks, rather than "harmless" short-of-war activities conducted in between hostilities.

While the necessity of public declaration within cyberwar is apparent, it is less clear what form such declarations should take. Despite dozens of states having been retroactively implicated in cyber operations across the world, state declarations claiming responsibility for cyberattacks are notably absent. Instead, the status quo has hitherto been for states to either wholly deny responsibility or remain entirely silent on the matter. Rarely, a state will acknowledge that it has "conducted cyber operations", [303] however it will refuse to divulge any further meaningful details regarding their efforts, such as the operation's objective or its methodology. While these vague declarations render it apparent that a state is engaged within the cyber domain, they do little to fulfill the intent of the public declaration principle as the information provided proves inadequate

[302] Orend, *War and Political Theory,* 89.
[303] The White House, Briefing Room, "The United States, Joined by Allies and Partners, Attributes Malicious Cyber Activity and Irresponsible State Behavior to the People's Republic of China," July 19, 2021. Accessed March 7, 2022. https://www.whitehouse.gov/briefing-room/statements-releases/2021/07/19/the-united-states-joined-by-allies-and-partners-attributes-malicious-cyber-activity-and-irresponsible-state-behavior-to-the-peoples-republic-of-china/.

for the purposes of accountability. In order to meaningfully fulfill the *jus ad bellum* principle of public declaration, I argue that any cyber declaration of responsibility must fulfill three specific criteria: 1) it must explicitly claim responsibility for a specific cyber operation, whether preceding its deployment, during the operation itself, or in its timely aftermath; 2) it must clearly state the justification for the deployment of cyber force; and 3) it must identify the intentions behind the use of force. In the absence of satisfactory state declarations of responsibility within the cyber context, we may look towards the efforts of non-state cyber actors to illustrate how this principle may reasonably, and responsibly, be fulfilled.

In stark contrast to state cyber operations, cyberattacks conducted by non-state entities are often accompanied by public declarations of responsibility. In some cases, these declarations serve a practical function; cyber criminals may claim responsibility to make their demands known and to issue further threats should their requests remain unmet, as is regularly the case with ransomware attacks.[304] However, declarations of responsibility made by less nefarious (or even potentially benevolent) cyber actors offer some insight into effective public declarations of responsibility within cyberspace. Following the launch of its invasion of Ukraine in early 2022, Russia quickly found itself assailed by DDoS attacks targeting numerous Russian government websites, including those of its Ministry of Defense and the Kremlin. These disruptive operations were followed by more invasive attacks against Russian streaming services and state media, replacing regular programming with pro-Ukrainian messages and raw footage of the conflict in Ukraine.[305] Responsibility for these attacks was claimed by the decentralized cyber collective Anonymous, as it publicly declared "war" against Russia in response to the state's unprompted and severe armed aggression against Ukraine.[306]

Although there are undoubtedly key structural differences between collectives like Anonymous and state governments, such as the degree of inherent anonymity individual

[304] Kari Paul, "Who's behind the Kaseya ransomware attack – and why is it so dangerous?" *The Guardian*, July 7, 2021. Accessed March 7, 2022. https://www.theguardian.com/technology/2021/jul/06/kaseya-ransomware-attack-explained-russia-hackers.
[305] Dan Milmo, "Anonymous: the hacker collective that has declared cyberwar on Russia," *The Guardian*, February 27, 2022. Accessed March 7, 2022. https://www.theguardian.com/world/2022/feb/27/anonymous-the-hacker-collective-that-has-declared-cyberwar-on-russia.
[306] Jennifer Medbury, and Paul Haskell-Dowland, "The hacker group Anonymous has waged a cyber war against Russia. How effective could they actually be?" *The Conversation*, February 28, 2022. Accessed March 7, 2022. https://theconversation.com/the-hacker-group-anonymous-has-waged-a-cyber-war-against-russia-how-effective-could-they-actually-be-178034.

members of the former enjoy, the public declaration made by Anonymous affiliates claiming responsibility for the cyberattacks targeting Russian interests offers a suitable template for public declaration within the cyber context more broadly. Firstly, Anonymous' declaration of interference with Russian interests dispels any uncertainty regarding the origin of its cyber operations, as the declaration immediately preceded the onslaught of DDoS attacks striking Russian government interests.[307] This clear claim of responsibility addresses the risk of mistaken attribution as it not only renders it clear that the collective has entered the conflict against Russia, but it also offers a clear link between the group and a specific set of cyber operations. This removes some of the ambiguity regarding which operations have been conducted by which parties. This is of particular importance on battlefields with multiple active groups; Russia's invasion of Ukraine in 2022 saw numerous cyber operators, such as the Belarus-based Cyber Partisans, as well as an ad-hoc unit of hackers affiliated with the Ukrainian Ministry of Defense, conducting independent cyber operations alongside Anonymous' own efforts.[308] Publicly claiming responsibility for specific operations works towards eliminating the attribution problem within cyberspace.

Secondly, the declaration serves to iterate the collective's justification for engaging in cyber conflict. Both the declaration of war and the entities targeted by the subsequent cyber operations are clearly framed as a response to Russian aggression, rather than being merely a random engagement or one conducted for financial gain. While public declarations made by decentralized collectives may admittedly inspire some uncertainty regarding whether the individual making the declaration is a proper representative authority for the whole, Anonymous' response to the Russian invasion of Ukraine nonetheless follows a behavioural pattern evidenced by the group's previous cyber operations.[309] Furthermore, public declarations made by governments will sidestep these concerns as states possess clearly defined representatives in the form of heads of state and other elected officials, ensuring that proper authority is readily discernible. Thirdly, announcements by Anonymous affiliates likewise offer further insight into the intentions and objectives of their ongoing cyber operations, as the group has asserted that it

[307] Milmo, "Anonymous: the hacker collective that has declared cyberwar on Russia."
[308] Joel Schectman, Christopher Bing, and James Pearson, "Ukrainian cyber resistance group targets Russian power grid, railways," *Reuters*, March 1, 2022. Accessed March 7, 2022. https://www.reuters.com/technology/ukrainian-cyber-resistance-group-targets-russian-power-grid-railways-2022-03-01/.
[309] Medbury and Haskell-Dowland, "The hacker group Anonymous has waged a cyber war against Russia. How effective could they actually be?"

seeks “to push information to the Russian people so they can be free of Putin’s state censorship machine”.[310] This announcement makes explicit the collective’s intention to counteract escalating Russian domestic censorship aimed at controlling the narrative of the conflict and pre-emptively reigning in domestic potential discontent. Accordingly, Anonymous’ declaration of its involvement in Russia represents a meaningful public declaration going beyond merely vague admissions of responsibility for cyber operations, and instead further identifying the set of cyberattacks they are responsible for, their justification for launching the cyber operations, and the intentions behind their efforts.

Insofar as the principle of public declaration by proper authority serves a vital transparency function within cyberwar, any state declaration of responsibility must likewise fulfill these three distinct criteria. Firstly, the state must claim responsibility for its own cyber operations to address the concerns posed by the attribution problem within cyberspace. Secondly, a declaration must render clear the justification for a cyber operation; by explicitly identifying the moral justification for a cyber use of force, external parties are better positioned to evaluate whether the principle of just cause has indeed been met. Thirdly, any public declaration of responsibility should be joined by a further statement identifying what the state seeks to accomplish with its deployment of cyber force. Not only would transparency of motive publicly fulfill the anti-consequentialist principle of right intention, but it would also bolster subsequent evaluations of consequentialist principles, such as proportionality and probability of success, considerations of which are in part contingent on what form a state’s wartime objectives take. Accordingly, tacit suggestions of involvement in vaguely defined cyber operations do not fulfill the cyber principle of public declaration. Rather, this principle demands more explicit declarations which clearly express a state’s responsibility, its justification, and its intentions. Only once these three requirements are met does a public declaration adequately address concerns of attributability, transparency and, ultimately, accountability.

[310] Vasco Cotovio and Mia Alberti, “Anonymous claims responsibility for "ongoing" hacking of Russian government sites,” *CNN*, February 26, 2022. Accessed March 7, 2022. https://www.cnn.com/europe/live-news/ukraine-russia-news-02-26-22/h_b7cb924af9172ecfce000603d4c8b5e9.

5.6 The Anti-Consequentialist Principles

Within the *jus ad bellum* of JWT, the principle of just cause is supplemented with additional considerations in the form of the principles of right intention and public declaration by the appropriate authority. In the case of a cyber-specific *jus ad bellum*, the principle of right intention necessarily constrains the tactics at a cyber-belligerent's disposal. On the one hand, a state's prioritization of cyber operations over kinetic alternatives may offer compelling evidence of its right intention insofar as such operations may reasonably achieve their objectives while inflicting comparatively minimal harms. Given that these operations may be transient and more readily reversible, they may represent meaningful efforts towards establishing Walzer's restoration plus in the post-war period. By contrast, cyber operations which seek to achieve their objectives by manipulating a state's common life and artificially exacerbating existing societal divisions are wholly impermissible. While these latter operations may realistically serve to end conflict between states, they amplify the likelihood of subsequent outbreaks of domestic unrest and civil strife. Ultimately, this serves to actively undermine the project of restoration plus by weakening the cohesiveness of the common life upon which the state is built.

Meanwhile, the principle of public declaration by the proper authority works towards ensuring that states remain accountable for their actions, both to the international community and to their own citizens. On a global stage, this principle ensures that any state of war between nations is made readily apparent, and any operation undertaken by one state against another is appropriately attributed, minimizing the likelihood of collateral damage, and ensuring that any necessary reparations or punitive measures for misconduct are made possible in the postwar period. Domestically, this principle works towards ensuring that a state's actions reflect the self-determined political values of its citizens, preventing a rogue state element from being able to operate in a purely self-interested manner against the best interests of its citizens. Within the context of cyberwar, this principle serves to address the pressing additional concerns posed by the attribution problem. Insofar as cyberattacks are inherently more covert than most of their kinetic counterparts, the principle of public declaration works towards preventing the mistaken attribution of cyberattacks which may needlessly escalate tense situations into full-blown conflict. Collectively, the anti-consequentialist principles of *jus ad bellum* offer an early evaluative framework which identifies the grounds upon which a state is justified in going to

war, as well as a pair of early responsibilities which necessarily accompany such a decision. In the next section, we will examine the remaining consequentialist principles of *jus ad bellum* and their applicability within the cyber domain so as to complete a fuller cyber *jus ad bellum* framework.

Chapter 6

The Consequentialist Principles

6.1 Completing the Cyber *Jus ad Bellum*

While the anti-consequentialist principles of just cause, right intention, and public declaration by proper authority establish necessary considerations to be made prior to declarations (and actions) of war, they alone are insufficient for establishing a comprehensive *jus ad bellum* framework. Despite the efforts of these principles to ensure a pre-war "fair set of rules that are binding on everyone", it would be short-sighted to exclude the consequential effects of war when distinguishing a just war from an unjust one.[311] We cannot claim a war is just solely because it fulfills the anti-consequentialist principles beforehand. It is well within the realm of possibility for a state to fulfill each of these three anti-consequentialist principles prior to waging war with the kinds of indiscriminate and brutal tactics that shock the moral conscience and dispel any illusion that the war it wages is just. We would be hard-pressed to assert that a combatant conducting a war in such a manner enjoys the same moral justification as one which fights with restraint, seeking to inflict no more harm than absolutely necessary for repelling an aggressor. Evidently, the consequences of conflict must necessarily be taken into account within any comprehensive evaluative framework. It has long been established that one needs to consider, in advance of resorting to war, what the likely results of one's war action might be.

To this end, the *jus ad bellum* framework of the Just War Theory (JWT) tradition supplements the preceding anti-consequentialist principles with three further consequentialist principles meant to account for war's consequences, both in actuality and reasonable forseeability: proportionality, last resort, and probability of success.[312] While the anti-consequentialist principles establish a set of criteria which determine whether a state is morally justified in resorting to war, these subsequent consequentialist principles add stipulations which further determine whether a state *should* go to war; that is, whether the state's decision to go to war will lead to good consequences, or whether the negative consequences of such a decision

[311] Brian Orend, *War and Political Theory*, (Cambridge, UK; Medford, MA: Polity, 2019), 92-93.
[312] Orend, *War and Political Theory*, 93.

irrevocably outweigh the positive ones.[313] In the interests of formulating a comprehensive *jus ad bellum* framework for cyber operations, this chapter will discuss each of these three consequentialist principles of *jus ad bellum* and examine how each principle translates into cyberspace.

6.2 Fighting Fire with Fire: The Principle of Proportionality

Although the possession of a just cause offers states a pretext for going to war, it cannot be relied upon as the sole justification for doing so. Even wars fulfilling each of the anti-consequentialist *jus ad bellum* criteria can be terrible affairs, as any substantial armed conflict between states often ensures that "the stage is set for a long series of terrible choices".[314] Military officers plan operations with the acute knowledge that certain types of collateral damage can prove unavoidable in the pursuit of their wartime objectives. On the ground, soldiers often find themselves pressed into critical decision between morally weighty options, such as determining whether to risk incurring greater losses amongst their ranks to protect the lives of civilians trapped in combat zones.[315] Further away from the battlefield, politicians deliberate grander wartime policies, seeking to satisfy foreign policy objectives while balancing domestic political and economic concerns; one of history's most notable terrible choices, the deployment of atomic weapons over the skies of Japan in 1945, was itself publicly justified by the US government as a preferable alternative to a costly land invasion.[316] Due to the inherent harms entailed by war, the principle of just cause is insufficient for determining whether a state *should* go to war. There must be further constraints which take these potential harms into account.

The principle of proportionality is one such constraint, working to ensure that any proposed response to an act of aggression is *commensurate to the harm incurred*.[317] Put simply, this principle asserts that force may only be employed when it represents a proportionate, measure, or reasonable response. Prior to any deployment of force, states must first evaluate

[313] Orend, *War and Political Theory*, 93.
[314] Gary D. Brown, "Proportionality and Just War," *Journal of Military Ethics* 2, no. 3 (November 2003): 171–85, https://doi.org/10.1080/15027570310000667. 175.
[315] Brown, "Proportionality and Just War," 175.
[316] Martin J. Sherwin, "Hiroshima as Politics and History." *The Journal of American History* 82, no. 3 (1995): 1085–93. https://doi.org/10.2307/2945113. 1085-1086.
[317] Orend, *War and Political Theory*, 86.

whether the "costs of the use of lethal force [are] outweighed by the value of what the lethal force is meant to accomplish, the military objectives of the use of force".[318] Within the specific context of the *jus ad bellum* of self-defense, Kilovaty conceptualizes proportionality as holding both quantitative and functional meanings. Quantitatively, proportionality demands that responses to acts of aggression do not exceed the scale and effects of the initial aggressive actions.[319] Supposing a state incurs a precision air strike on a military facility by a neighbouring nation, it would have just cause for responding with force. In the interests of responding in a deliberately heavy-handed manner so as to deter any further potential for aggression, the victim state may respond with large-scale shelling of the aggressor's capital city. This would represent a grievous violation of the principle of proportionality insofar as the scale and effects of the retaliatory action far surpass those of the initial act of aggression. Insofar as the retaliation would result in a far greater degree of harm, it is readily recognizable as a disproportionate response. A more proportionate response could see the victim state deploying precision air strikes in kind, targeting solely the airbase from which the aggressor had launched their attacks. This alternative would prove quantitatively proportional on account of bringing about a degree of harm similar to that incurred by the original act of aggression.

Functionally, the principle of proportionality demands that any use of force must be *proportionate to the objective it seeks to accomplish*.[320] This element of proportionality seeks to ensure that the ends ultimately justify the means. To illustrate, suppose a state's neighbour has been conducting regular precision drone strikes against military equipment stockpiles near their shared border. A large-scale ground invasion would undoubtedly prove quite effective at halting an aggressor's ability to continue conducting drone operations, however it would be controversial to claim that a full military occupation represents a proportionate response to localized drone strikes that have hitherto failed to result in bodily harm or death. On the other hand, a special forces sabotage operation against the specific airfield housing the aggressor's drone capabilities would offer a more proportionate means of taking an aggressor's drone capabilities offline. The principle of proportionality limits the kinds of responses states may

[318] Eric Boylan, "Applying the Law of Proportionality to Cyber Conflict: Suggestions for Practitioners," *Vanderbilt Journal of Transnational Law*, vol. 50, no. 1 (January 2017): 217-246. 227.
[319] Ido Kilovaty, "Cyber Warfare and the Jus Ad Bellum Challenges: Evaluation in the Light of the Tallinn Manual on the International Law Applicable to Cyber Warfare," *National Security Law Brief* 5, no. 5 (2014): 90-124. 103.
[320] Kilovaty, "Cyber Warfare and the Jus Ad Bellum Challenges," 103.

justifiably deploy by restricting those options which go well beyond what is necessary for achieving their objectives. As a result, the principle of proportionality works towards mitigating the risk of escalation which accompanies overly zealous responses to aggression. Insofar as a response is proportionate to the aggression incurred, it is less likely to itself be responded to with disproportionate force.

The principle of proportionality may have different demands across varying kinds of conflict, as Brown suggests that not all just wars may be possible to fight in a proportionate way. Brown argues that war is more readily waged proportionally when it is directed against an identifiable, centralized target, such as a government.[321] These kinds of entities possess easily recognizable infrastructure and regularly have some sort of military apparatus; accordingly, legitimate targets, such as soldiers and government installations, are more clearly designated within conflicts involving these entities. This renders it easier to wage war lawfully. By contrast, wars waged against decentralized entities, such as terrorist organizations, present a greater challenge for the proportionality principle. As evidenced by the Global War on Terror (GWOT), asymmetric warfare regularly involves combatants embedded within the public so as to improve their fighting effectiveness against a superior conventional force. These combatants not only prove difficult to reasonably identify, but they are also difficult to fight within the constraints of proportionality as their embedded nature makes targeting them likely to result in disproportionate collateral damage to civilian lives and civic infrastructure. While asymmetric conflicts such as the GWOT might fulfill the functional component of proportionality, insofar as war is a proportionate response to acts of extreme aggression, these conflicts face a serious hurdle in the form of Kilovaty's quantitative element of proportionality; the elevated likelihood of collateral damage to civilian interests in the pursuit of these wars places a greater strain on a belligerent's ability to wage such wars proportionately. In the event of irregular conflict, the principle of proportionality necessarily places greater constraints on how such wars may be justly pursued.

Just as the principle of right intention necessarily extends into the *jus in bello* stage of conflict, so too does the principle of proportionality demand further attention as war unfolds. The battlefield is not a static setting. While belligerents enter war with a set series of objectives, they must necessarily adapt to the objectives and tactics of their opponents as battles are won and lost.

[321] Brown, "Proportionality and Just War," 175.

The capabilities of combatants likewise change radically over the course of conflict as military losses diminish their ability to conduct certain types of operations, while emergency arms production may bolster other warfighting capabilities. Further complicating matters, the development of nascent technologies and tactics may emerge as niche strategic problems arise within specific conflicts, further altering the landscape of any given war. As wartime strategy adapts to these changes, so too does the principle of proportionality require constant revisiting and adjustment. Artillery barrages against an aggressor's fuel infrastructure may prove initially proportionate insofar as its ability to wage war relies significantly upon the strength of its armoured military units, despite the detrimental effect these attacks may have on civilians reliant on the same infrastructure for heating and mobility. However, as the conflict continues and the aggressor suffers extensive losses to its armoured units, these previously proportionate attacks against oil and gas infrastructure may begin to be regarded as disproportionate once risk of further armoured aggression dwindles. In this case, the principle of proportionality would motivate the gradual introduction of greater restraint as the war goes on to mitigate collateral harms inflicted against illegitimate targets of war. It is insufficient to assert that conduct *within* war is proportionate simply because it was deemed proportionate *prior* to the outset of hostilities. Rather, the "calculus of proportionality … is a continuing one", requiring the revisitation of pre-war deliberations of proportionality throughout conflict to ensure that a belligerent's conduct remains proportionate throughout.[322]

The principle of proportionality works to temper the just cause principle by ensuring that any potential response by a state to an incurred act of aggression is itself proportionate. A state being justified in going to war does not mean that it *should* do so. In other words, even if one has the right to go to war, it may still be inadvisable to proceed, depending on the predicted costs of exercising that entitlement. Even in the best of cases, war is inherently destructive. Any forceful response to aggression must be predicated by calculations determining whether the benefits of such a response outweigh the costs of a forceful course of action. According to the principle of proportionality, the use of responsive force is justifiable if 1) it is proportionate to the harm that has been incurred; and 2) it is proportionate to the end objective it seeks to accomplish. Responding to an act of aggression with disproportionate force may prove successful at

[322] Brown, "Proportionality and Just War," 176.

preventing further acts of aggression, however such a response would inflict a greater degree of harm than has been incurred. In practice, the principle of proportionality works to mitigate the proliferation of undue harms and the likelihood of escalations in hostilities which may accompany disproportionately forceful responses.

6.3 Proportionality Constraints on Cyberwar Responses

The growing emergence of cyber operations as viable alternatives to kinetic operations introduces novel considerations for the proportionality principle and its translatability into cyberspace. Notably, existing LOAC conceptions of the proportionality principle do not offer any restrictions regarding *kind* of force and proportionality. Rule 72 of the *Tallinn Manual* suggests that "there is no requirement that the defensive force be of the same nature as that constituting the armed attack", noting that some adversaries may prove digitally resilient while remaining vulnerable to conventional responses, motivating responses along a different axis than the initial act of aggression.[323] Accordingly, a state incurring a kinetic act of aggression is not limited to responding to the attack with kinetic force; it may instead opt to respond with the deployment of cyberweapons. Likewise, the opposite holds true in theory, as the proportionality principle of the LOAC does not preclude a state's ability to respond to a severe cyberattack with kinetic force. Provided the response is of a similar scale and effect to that of the incurred aggression, a state is free to deliver its response either digitally or kinetically. While the LOAC approaches may treat kinetic and cyber operations as largely interchangeable with regards to the principle of proportionality, I argue that the idiosyncrasies of cyber operations render them more likely to fulfill the principle of proportionality than conventional operations. This holds particularly true for instances within which an incurred cyberattack is deemed severe enough to motivate a forceful response.

The unique methodology and capabilities of cyberwar, irreplicable by conventional alternatives, has led to assertions that cyberwar represents a form of "ideal war".[324] An ideal war

[323] Michael N. Schmitt, and NATO Cooperative Cyber Defence Centre of Excellence, eds. *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations*. Second edition. Cambridge, United Kingdom; New York, NY, USA: Cambridge University Press, 2017. 349-350.
[324] Ryan Jenkins, "Cyberwarfare as Ideal War," in *Binary Bullets: The Ethics of Cyberwarfare*, eds. Fritz Allhoff, Adam Henschke, & Bradley Jay Strawser (New York, NY: Oxford University Press, 2016), 89.

is construed as an armed conflict within which "civilian casualties were minimal or nonexistent and where acts of violence perfectly discriminated between combatants and noncombatants".[325] Problematically, this form of ideal war has hitherto been unattainable. Conventionally waged wars regularly result in a plethora of widespread harms, from loss of civilian life to mass displacement of refugees, as conflict devastates municipal infrastructure and renders cities uninhabitable. Even the deployment of discriminate force, such as guided bombs and unmanned drone strikes, often results in damage extending beyond the immediate target of force; while being touted as a discriminate means of deploying force with greater precision than conventional munitions, drone strikes have nonetheless resulted in hundreds of collateral civilian deaths over the course of their operational history.[326] While such an ideal war may be possible, it is difficult to see it being achieved with kinetic weapons and tactics, despite continuous efforts to make such weapons more discriminate.

However, the emergence of cyberwarfare has led to renewed optimism that ideal war may be a realistic possibility. In contrast to the harms inherent to conventional war, Jenkins asserts that cyberwar harms can be "perfectly discriminate and totally reversible".[327] To illustrate the more discriminate nature of cyberweapons, Jenkins points towards the deployment of the Stuxnet worm which proved capable of sabotaging Iranian centrifuges through selectively infecting only those computers running a specific kind of software designed to engage with programmable logic controllers.[328] In working its way towards the relevant centrifuge control systems, the worm only spread to systems which met this criterion, suggesting the worm was designed to discriminate between potential targets and non-target computer systems. Ideally, such a cyberweapon would be designed to entirely avoid infecting the computer systems of illegitimate targets, such as those serving a critical role within civilian infrastructure. While Stuxnet infected thousands of computers beyond those immediately associated with Iran's nuclear enrichment program, Jenkins argues that worm nonetheless remained discriminate by virtue of it remaining inert on non-target computers and harming only the computer systems of its target.[329] Should

[325] Jenkins, "Cyberwarfare as Ideal War," 89.
[326] Scott Shane, "*Drone Strikes Reveal Uncomfortable Truth: U.S. Is Often Unsure About Who Will Die*," *The New York Times*, April 23, 2015. Accessed March 15, 2022. https://www.nytimes.com/2015/04/24/world/asia/drone-strikes-reveal-uncomfortable-truth-us-is-often-unsure-about-who-will-die.html.
[327] Jenkins, "Cyberwarfare as Ideal War," 96.
[328] Jenkins, "Cyberwarfare as Ideal War," 97.
[329] Jenkins, "Cyberwarfare as Ideal War," 97.

cyberweapons be designed to inflict harm solely against legitimate targets, they could safely address Brown's concerns regarding the inability to proportionately conduct certain kinds of conflict. Insofar as cyberweapons prove significantly more discriminate than their conventional counterparts, they represent a use of force which may be deployed proportionately against precise and intended targets without the corresponding risk of inflicting disproportionate harms against civilians in the process.

Beyond the comparatively discriminate nature of cyberweapons, Jenkins points towards two further boons for the ability of cyber operations to fulfill the principle of proportionality. The first of these is that, unlike conventional action, cyberweapons are "maximally proportionate" insofar as they can achieve their objectives "while inflicting totally reversible harms".[330] Disruptive cyberattacks offer operators a means of achieving their objectives without necessarily inflicting lasting harms. For example, a cryptographic attack enables a cyber operator to encrypt key programs on a target's computer systems, potentially locking the target out of various programs integral to the state's day-to-day functions until the attacker decrypts the program.[331] Alternatively, a resource-deception attack may be employed to mask the normal functionality of a target's systems by generating falsified error reports which may undermine a target's trust in its own computer systems and hinder its ability to operate normally.[332] These kinds of operations can be remotely ceased with a few keystrokes once an attacker's demands are met or their objective has been achieved. As such, their effects are reversible in mere moments, rather than requiring the weeks or months of reconstruction efforts which typically follow in the wake of kinetic operations; this allows for the more immediate restoration of pre-war functionality than is possible with kinetic alternatives. Insofar as a use of kinetic force is deemed a proportionate response to incurred aggression, it stands to reason that a cyber approach resulting in fewer collateral harms and less permanent (or long-term) damage would represent a *more proportionate* alternative.

Secondly, the flexibility of cyberweapons may enable them to be continuously proportional responses throughout a conflict, bolstering their ability to address *in bello* concerns

---

[330] Jenkins, "Cyberwarfare as Ideal War," 98.
[331] Neil Rowe, *Towards Reversible Cyberattacks*, Reading: Academic Conferences International Limited, 2010. http://search.proquest.com.proxy.lib.uwaterloo.ca/conference-papers-proceedings/towards-reversible-cyberattacks/docview/869507120/se-2?accountid=14906. 263.
[332] Rowe, *Towards Reversible Cyberattacks*, 265.

of proportionality. While it is possible to deploy cyberweapons which operate autonomously in accordance with pre-set instructions, Jenkins suggests that other cyberweapons "could be programmed to communicate constantly with command and control servers elsewhere".[333] By designing cyberweapons to maintain a constant line of communication, a cyberweapon may offer crucial real-time information to the operators who had deployed it. Information such as a computer worm's spread across target- and non-target computer systems, up-to-date assessments of a target's capabilities, and reports of the current effects a cyberweapon has had on target computer systems all work towards providing cyber operators with a more accurate evaluation of the virtual battlefield from which they may make more informed and on-going assessments of proportionality. Likewise, communication lines between the operators and the weapon allows the former to maintain constant control of the latter, rendering it possible for the harms of a cyberweapon to be "redirected, terminated, or reversed if need be".[334] Insofar as cyber operators remain capable of making minute adjustments to the scope and effects of their cyberweapons even after their deployment, it is possible to ensure that cyberweapons remain proportionate uses of force throughout a conflict.

While these features of cyberweapons support the conclusion that cyber weapons are generally more likely than kinetic alternatives to fulfill both the quantitative and functional elements of proportionality, it is worth addressing the interchangeability of kinetic and cyber responses within the LOAC conception of proportionality to discern whether kinetic responses to cyberattacks may ever be warranted in accordance with this principle. The *Tallinn Manual* Experts suggest there are no immediate constraints on the *kind* of response a targeted state may employ in the event of aggression rising to the level of an armed attack; both cyber and kinetic responses are deemed potentially permissible provided they result in the same scale and effects of the incurred act of aggression. A missile attack on a military weapons storage facility resulting in significant physical destruction and bodily injury may motivate a retaliatory strike targeting the aggressor's missile batteries; whether such an attack manifests as a series of drone strikes against the aggressor's missile facilities or the remote detonation of warheads made possible by a penetrative cyberattack is ultimately inconsequential provided that the scale and effects are

[333] Jenkins, "Cyberwarfare as Ideal War," 97.
[334] Jenkins, "Cyberwarfare as Ideal War," 98.

proportionate to the purpose of repelling further aggression.[335] While a cyberattack may represent a *more proportionate* response, given its ability to achieve similar objectives with less corresponding harm, a kinetic response would nonetheless fulfill the proportionality principle in this case by virtue of being *similarly proportionate*. This would likewise hold true if the consequences of the initial aggressive attack against the weapons storage facility were conducted by purely digital means. Insofar as the scale and effects of the attack are similar across the kinetic-cyber divide, both kinetic and cyber responses represent potentially proportionate answers in accordance with the proportionality principle of a cyber *jus ad bellum*.

However, the calculus of proportionality becomes more difficult when a cyberattack results in disanalogous harms. Suppose State A, motivated by a heated territorial dispute, launches a cyber operation against State B to force it to rescind its territorial claims. Rather than seeking to cause physical destruction, State A's cyber operation may instead target a variety of State B's government computer systems to delete swathes of crucial data, backing up the destroyed information on servers within State A for restoration once State B capitulates to its demands. State A's cyber operation may result in heavy disruptions to State B's state-run healthcare services, for example, as vital patient information is wiped from its servers. Similarly, state-affiliated rail networks may likewise be rendered inoperable should such an attack lock State B out of its railway control systems. It may also seek to manipulate domestic election results to prop up a political candidate which has publicly recognized State A's claim to the territory. Such an operation may feasibly bring State B's regular functioning to a halt while likewise posing a threat to State B's common life, all without resulting in physical harm or destruction. The effects of this cyberattack would prove wholly reversible should State A choose to call off its operation. In this case, State A's cyber operation fulfills the just cause principle for State B insofar as it represents an aggressive violation of State B's sovereignty designed to advance State A's political objectives.

Despite State A's attacks qualifying as an act of aggression, it is difficult to assert that a kinetic response by State B would be proportionate in this case. The scale and effects of harm caused by State A's cyber operation are not readily replicable through conventional means; it would be difficult to conceive of a kinetic operation which could result in the same degree of

[335] *Tallinn 2.0*, 349.

disruptiveness without a substantial increase in physical damage and bodily harm. Likewise, the disruption caused by these concerted cyberattacks appears materially different to the sort which may be caused by conventional alternatives insofar as the incurred damages are intangible and readily reversible. As was the case with earlier discussions regarding just cause, the disanalogous nature of non-physically destructive cyberattacks renders proportionality comparisons with kinetic operations quite difficult due to the comparatively little common ground to be found between the two approaches.

I argue that the disanalogous nature of certain kinds of cyberattacks necessarily precludes the ability of states to respond kinetically due to two specific difficulties arising in comparisons between cyber- and kinetic operations. Firstly, insofar as the principle of proportionality relies on comparisons of scale and effects between acts of aggression and their responses, the irreplicability of non-conventional harms by conventional means renders uncontroversial comparisons of harms impossible. A set of harms readily achieved by both kinetic and conventional means makes the proportionality calculus straightforward; should the scale and effects of an incurred attack be replicable by either kinetic or cyber means, the responding state may justifiably deploy either form of response while still adhering to the proportionality principle. If a cyberattack against rail control systems results in the derailment of a freight train and the destruction of a section of tracks, the victim state may deploy a kinetic response destroying a freight train and severing a section of rail; in both attack and response, the scale and effects are recognizably similar and, as such, appear uncontroversially proportionate. By contrast, a cyberattack resulting in the deletion of swathes of critical government data results in harms not readily replicable by conventional means. In the absence of the possibility of responding with like-to-like harms, the proportionality principle may only be met if the distinct harms inflicted by a response are evaluated as being proportionate to the distinct harms incurred by the initial attack of aggression.

This leads to the second difficulty with deploying kinetic responses to cyber acts of aggression. As evidenced by the inability of the *Tallinn Manual* Experts to form a consensus regarding the use-of-force classification of disanalogous cyberattacks,[336] gauging the relative severity of non-physically destructive cyberattacks is a difficult process and one fraught with

[336] *Tallinn 2.0*, 342.

controversy. While harms such as physical destruction lend themselves to straightforward evaluations of severity irrespective of their origin, evaluating how a non-physical harm compares to the harms of a potential physical response proves more challenging. Would a cross-border sabotage raid against air defense systems be seen as resulting in proportionate harms to cyberattacks encrypting vital military communications data? Would cyber efforts to jam naval navigation systems for naval fleets warrant responding with a physical attack against the communications infrastructure rendering such efforts possible? In both cases, assertions made regarding the relative severity of the harms incurred are likely to remain controversial due to the dissimilarity of the kinds of harms being wrought. Insofar as the harms caused by non-physically destructive cyberattacks are disanalogous to those inflicted by kinetic attacks, we lack an appropriate point of reference from which to reach uncontroversial assessments of their severity. As a result, achieving consensus regarding comparisons between disanalogous harms is likely to prove wholly impossible, especially between states currently locked in conflict. The result is general uncertainty regarding how a kinetic response may be weighed against a cyber use of force. This uncertainty increases the risk of a mistaken evaluation of proportionality which potentially triggers an escalation of hostilities.

In the interests of preserving the pragmatic function of the principle of proportionality, I argue that the interchangeability of kinetic and cyber responses to cyberattacks applies only in cases within which the initial cyber aggression results in harms analogous to those of conventional operations. The international community has a wealth of experience evaluating the kinds of harms arising from kinetic conflict. This has led to a degree of familiarity with comparing the severity of kinetic operations, even in cases within which the specific effects of two operations are different. As a result, the severity of cyberattacks resulting in kinetic-analogous harms may be reasonably evaluated with reference to this history of experience as a common ground; this opens the possibility of responding to such cyberattacks with kinetic measures. However, this kind of familiarity is absent in the case of comparing cyber-specific harms with kinetic ones, undermining the common ground upon which agreements regarding the proportionality of kinetic responses to cyber harms may be made. Without a common reference point, a kinetic response may be perceived by one state as a proportionate response to incurred cyber aggression, while simultaneously being received by other members of the international community as wholly disproportionate. The relative uncertainty of evaluations regarding the

severity of cyber action is magnified further by general disagreements pertaining to how concepts such as ownership and territory translate into the digital domain; some states may perceive some invasive cyber operations as particularly severe infractions due to unshared interpretations of territory within cyberspace.[337] Insofar as these controversies are unlikely to be resolved soon, there is pragmatic value in centering the principle of proportionality on more secure grounds.

While familiarity with the consequences of kinetic action motivates the permissibility of kinetic responses to cyberattacks resulting in similar kinds of harms, cyber operations resulting in novel cyber harms should only be responded to with cyberattacks in kind. Limiting permissible responses to disanalogous cyberattacks to solely those operations capable of inflicting similar harms restricts the ability of states to respond to incurred harm by inflicting harm of a radically different sort. Insofar as the harms of both the act of aggression and the victim's response are of a similar type, evaluative efforts would avoid the difficult hurdle of comparing harms across the physical-digital divide. This approach would serve to reduce the risk of radically divergent evaluations of proportionality which may inadvertently trigger escalations of force between states. Furthermore, the less inherently destructive nature of cyber operations mitigates the likelihood of a disproportionate cyber response leading to the same kind of escalations. Unlike kinetic attacks, cyber operations are often more transient and more readily reversible; while the risk of disproportionate cyber responses exists, the consequences of a cyber overreach pale in comparison to those of a disproportionate kinetic response.

As a result, I argue that the principle of proportionality within cyberspace ought to be bolstered with pragmatic considerations to mitigate the risk of controversial evaluations of severity and the correspondingly disproportionate responses they may entail. While cyberattacks may hypothetically warrant kinetic responses, I argue that kinetic responses are justified if and only if the initial cyberattack itself results in kinetic-analogous harms, such as physical destruction or bodily injury. Insofar as these harms are familiar, they are more readily evaluable through existing frameworks of proportionality. In contrast, novel cyber harms present disparate kinds of damages which fail to offer immediate points of comparison to conventional harms. In the absence of common ground, such harms should be considered fundamentally *distinct* from

[337] George R. Lucas Jr., "Emerging Norms for Cyber Warfare," In *Binary Bullets: The Ethics of Cyberwarfare*, eds. Fritz Allhoff, Adam Henschke, & Bradley Jay Strawser (New York, NY: Oxford University Press, 2016), 20.

kinetic harms to address concerns of radically divergent proportionality comparisons between cyber- and kinetic harms. Insofar as cyberattacks can achieve similar objectives to conventional operations, albeit with comparatively fewer harms, cyberattacks may represent proportionate responses to kinetic attacks. However, the difficulties surrounding the evaluation of cyber-specific harms, and the risks associated with "getting it wrong", should necessarily restrict the ability of states to respond to novel cyberattacks with kinetic force.

6.4 The Principle of Last Resort

Although uses of force offer undeniably effective means of deterring aggression, they do not do so without causing further harms of their own in the process. Consider State A's attempts to deter State B's cross-border air incursions which have just resulted in the partial destruction of the airbase home to State A's drone surveillance capabilities. Insofar as State B's attacks are unprovoked uses of force, State A has just cause for its own deployment of force. Likewise, the government of State A may publicly declare its intention to intervene with the sole objective of hampering state B's air capabilities, as well as publicly committing to operating in accordance with the existing LOAC; in so doing, State A would fulfill both principles of right intention and public declaration. Finally, State A may opt to fight fire with fire, retaliating against State B's air raids with airstrikes of its own, resulting in similar scale and effects to State B's aggression and thereby satisfying the principle of proportionality. Despite these *jus ad bellum* principles having been satisfied, State A's response nonetheless results in the proliferation of *further harm*. While a response on State A's behalf might prove necessary (assuming State B's aggression is likely to continue or intensify in the absence of a deterrent), we would be hard-pressed to assert that State A's actions would be *equally morally justifiable* to an alternative which curtails State B's aggression while resulting in *less corresponding harm*. In cognizance of this, the principle of last resort emerges as a further constraint on how liberal states may be in their deployment of forceful countermeasures to incurred acts of aggression.

Simply put, the principle of last resort permits responsive uses of force if and only if non-forceful alternatives have already been exhausted.[338] Despite the effectiveness of forceful

[338] Orend, *War and Political Theory*, 87.

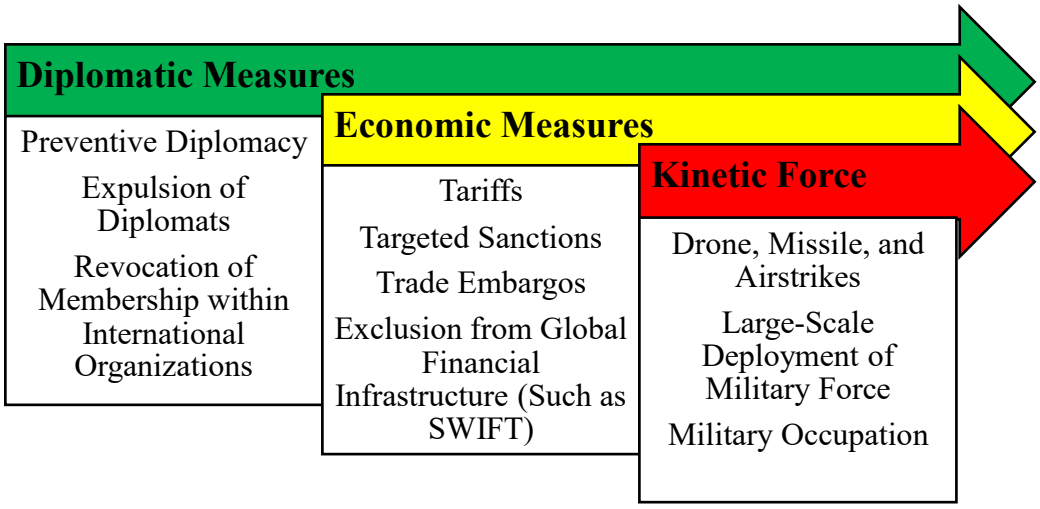
countermeasures, force is not a one-size-fits-all necessary answer to all forms of interstate aggression. Rather than immediately resorting to force, states should explore diplomatic alternatives, including strategies of symbolic denouncements, such as shutting down embassies within the aggressor's borders or expelling its diplomats, to more drastic measures, such as the suspension of an aggressor's membership within international organizations; this latter option was on display in the G8's decision to forgo continuing dialogue with the Russian Federation following its annexation of Crimea in 2014.[339] Should measures of diplomacy prove unfruitful, states may escalate the severity of their response while nonetheless avoiding a resort to force. Economic sanctions offer an intermediary measure between diplomatic and kinetic responses, allowing for greater coercive pressure while remaining less likely thank kinetic responses to trigger an escalation of hostilities. The potency of economic measures was evidenced in early 2022 as NATO-aligned countries sought to economically isolate Russia in response to its invasion of Ukraine; in hopes of bringing Russian aggression to a halt, states sanctioned a range of Russian-affiliated targets, from Russian domestic manufacturing to Russian economic institutions, and even oligarch-owned sports teams abroad.[340] Despite assertions by the Russian government that concerted economic sanctions of such a scale would be considered an "act of war",[341] the lack of immediate escalation of hostilities against Western states suggests that the Russian government nevertheless treats economic warfare as materially different to actual kinetic involvement. Accordingly, even severe economic measures offer alternative means of response while still falling short of a use of kinetic force.

The efforts by the US-led coalition to motivate an Iraqi withdrawal from Kuwait in the early 1990s offer an example of how the principle of last resort manifests in practice, and how the principle may be satisfied by a responding state or organization. In seeking to respond to Iraq's aggressive invasion of Kuwait in 1990, the coalition did not immediately resort to kinetic war as the default approach to ousting the Iraqi military. Rather, initial efforts manifested at the

[339] Julian Borger and Nicholas Watt, "G7 Countries Snub Putin and Refuse to Attend Planned G8 Summit in Russia," *The Guardian*, March 24, 2014. Accessed March 21, 2022. https://www.theguardian.com/world/2014/mar/24/g7-countries-snub-putin-refuse-attend-g8-summit-russia.
[340] David Leonhardt and Ian Prasad Philbrick, "*Economic War*," *The New York Times*, March 11, 2022. Accessed March 21, 2022. https://www.nytimes.com/2022/03/11/briefing/economic-war-sanctions-russia.html.
[341] Toby Helm, Luke Harding, Daniel Boffey, and Julian Borger, "Defiant Putin Warns the West: Your Sanctions are Akin to an Act of War," *The Guardian*, March 5, 2022. Accessed March 21, 2022. https://www.theguardian.com/world/2022/mar/05/defiant-putin-warns-the-west-your-sanctions-are-akin-to-an-act-of-war.

diplomatic level as coalition governments first sought a diplomatic solution to the problem.[342] Once diplomatic efforts were rebuffed, the focus turned towards the implementation of economic sanctions meant to incentivize an Iraqi retreat from Kuwaiti territory.[343] As economic sanctions proved similarly ineffective, the coalition's focus then turned towards conventional military alternatives. Instead of immediately launching a military offensive, the coalition first sought to telegraph its intention to invade with military buildups along Saudi Arabia's borders in hopes of intimidating the Iraqi government towards the negotiating table.[344] Only once this show of force likewise failed to convince the Iraqi leadership to pursue non-violent resolutions did the coalition resort to the deployment of conventional military force in 1991. The coalition's gradually escalating response strategy fulfilled the principle of last resort insofar as a spectrum of non-violent alternatives were first explored and deemed ineffective prior to the usage of inherently harmful kinetic action. This kind of approach ensures that kinetic countermeasures are deployed *only when absolutely necessary*.

Figure 1. Traditional Response Escalation



Although the principle of last resort demands that non-forceful alternatives be first exhausted before resorting to more harmful physical options, it should not be construed as necessitating a literal attempt of every conceivable short-of-war measure prior to morally justifying a forceful response. As Orend suggests, there is "always *something else* which could

[342] Orend, *War and Political Theory*, 88.
[343] Orend, *War and Political Theory*, 88.
[344] Orend, *War and Political Theory*, 88.

be tried".[345] The range of potential state responses to aggression is limited solely by the imagination; to expect a state to attempt every possible non-forceful option in hopes of achieving a peaceful resolution is unreasonable. Likewise, even amongst the options that have hitherto proven potentially successful alternatives (such as diplomatic and economic sanctions), there is a certain point at which repeated attempts at unsuccessful solutions may be safely omitted and a state is permitted to escalate the force of its response. Deliberations regarding which options should be reasonably exhausted by a state prior to turning towards "last resort" measures are contingent upon a myriad of factors to be determined on a case-by-case basis. Everything from the severity of the force incurred to the capabilities of a responding state may influence whether a certain resort to force ultimately fulfills the principle of last resort.[346]

For example, a state with a predominantly insular economy is unlikely to be hard hit by economic sanctions, potentially precluding the deployment of economic sanctions as a necessary step prior to exploring more forceful alternatives. In some cases, the broader context of conflict may lead to one state's mobilization for war being deemed as fulfilling the last resort criterion, while the same may not hold true if another state were in the same situation; a small state with little economic or political weight would be more readily justified in responding to military aggression with force in kind than a wealthier state which enjoys reasonable chances of deterring further aggression through leveraging its economic or political strength. This may be supplemented with further considerations regarding whether a state's inaction or failed non-physical responses might themselves result in "unreasonable, unacceptable, unjust consequences".[347] Supposing a state has committed to a campaign of rapid expansionistic aggression designed to annex its immediate neighbours, the threatened states may forgo longer-term economic and political responses in favour of an immediate military response as not doing so would be likely to result in unacceptable harms in the form of the invasion of the state. In these cases, the severity and imminence of the threat posed by the aggressor permits a quicker escalation to the level of kinetic retaliation.

The principle of last resort does not seek to prohibit forceful responses to aggression. Rather it works to ensure that the usage of force is *necessary*. Despite placing constraints on the

[345] Orend, *War and Political Theory*, 87.
[346] Orend, *War and Political Theory*, 87.
[347] Orend, *War and Political Theory*, 87.

permissibility of force as a response to aggression, the principle does not require that a state literally exhaust every possible alternative to force before justifying the deployment of a forceful option. A history of observable state policy and behaviours enables states to make rational conclusions regarding the effectiveness of certain kinds of short-of-force measures prior to attempting them. Insofar as a short-of-force alternative enjoys reasonable chances of success, it ought to be attempted prior to an escalation towards more forceful alternatives. On the other hand, alternatives with merely fringe chances of success may be safely omitted from a state's response strategy. Accordingly, the principle of last resort supplements the preceding *jus ad bellum* criteria by restricting deployments of force wherever less harmful short-of-force measures would suffice.

6.5 A Measure of Later Resort

Historic experience negotiating conventional interstate conflicts has led to the emergence of a linear escalatory progression of potential state responses to aggression. The use of diplomatic tools towards conflict resolution is widely regarded as the first step and least-severe option at a state's disposal. Once these efforts fail, the deployment of economic sanctions represents a natural escalation in response strategy; while economic sanctions can range dramatically in scope and effects, they are generally acknowledged by the international community as a more severe response than purely diplomatic efforts. In the same vein, current norms of interstate conflict resolution suggest a further severity divide between economic sanctions and the deployment of force. Due to the risks inherent to forceful responses, both in terms of their immediately harmful consequences and the corresponding risk of such measures working to escalate conflict, these kinds of responses are uncontroversially treated as a far more severe option than diplomatic and economic alternatives. Although states occasionally employ rhetoric suggesting that economic sanctions may be regarded as tantamount to a use of force, current practice suggests consensus amongst the international community that economic measures remain a tier below deployments of kinetic force in the severity of response hierarchy.

The idiosyncratic nature of cyberweapons complicates this relatively straightforward conception of response escalation due to both the novel methodologies of cyberweapons and

their current deployment practices. Firstly, the novel means by which cyberweapons achieve their objectives complicates their placement on the response hierarchy. As evidenced within the preceding chapters, cyberweapons have proven fully capable of supplanting kinetic weapons in certain capacities, resulting in similar scale and effects as would be expected of kinetic means. While the analogous harms wrought by these kinds of cyberweapons would necessarily mark them as a more severe response than economic sanctions, there are nonetheless unignorable differences between how cyber- and kinetic weapons achieve their objectives which render them less outwardly severe than kinetic options.

Kinetic weapons inescapably entail physical destruction, often to a degree which spreads harm beyond its immediate target. While advances in kinetic weapons technologies have led to the development of ever more precise weaponry, the novel methodologies of cyberattacks ensure that cyber options are nonetheless "likely to cause fewer civilian casualties than even the most carefully designed and executed kinetic attack".[348] Cyberweapons may be designed from the ground up to be "carefully crafted and targeted to affect only specific systems and organizations, greatly reducing undesired collateral effects", rendering them more likely to be perfectly discriminate.[349] Beyond the mitigation of potential harms inflicted against civilians within proximity of a target, cyberweapons pose less of a risk to military personnel of the responding state itself, as means of cyberweapon delivery typically lack the same element of risk which accompanies the deployment of conventional force; even if target resistance is unlikely, latent risks such as equipment malfunctions may inflict unforeseen losses on a responding state.[350] Provided that a cyber alternative may achieve its objective with a lower degree of expected harm, it is difficult to assert that the kinetic option and the cyber alternative represent equally severe responses; the latter is less likely to result in the scale and effects of harms which add significant moral weight to decisions to deploy the former.

Secondly, the current practices of the international community with regards to the deployment of cyberweapons suggests they are widely regarded to be more acceptable than

---

[348] James M. Acton, "Cyber Weapons and Precision-Guided Munitions," in *Understanding Cyber Conflict: Fourteen Analogies.* Eds. Perkovich, George, and Ariel E. Levite. Washington: Georgetown University Press, 2017. 45-60. muse.jhu.edu/book/62546. 45.

[349] D. Raymond, G. Conti, T. Cross and R. Fanelli, "A control measure framework to limit collateral damage and propagation of cyber weapons," *2013 5th International Conference on Cyber Conflict (CYCON 2013)*, 2013, pp. 1-16. 13.

[350] Acton, "Cyber Weapons and Precision-Guided Munitions," 45.

kinetic force. The ability of cyberweapons to be deployed discretely and discriminately has led to states treading them "*not* as a last resort, but rather as a first-strike capability",[351] a means by which to advance foreign policy objectives more directly than achievable with political approaches, while still avoiding the international backlash which regularly accompanies the deployment of kinetic force. State-backed cyberoperations have hitherto made extensive use of the grey area of moderation left unaddressed by the conventionally-minded LOAC. In the absence of the governance of binding international laws, we may glean the relative severity of cyberattacks by virtue of how they are received by the international community. The DDoS attacks against Estonia in 2007 were regarded as more severe than economic sanctions, insofar as the Estonian government had pushed for the attacks to be seen as an Article 5 violation by its NATO allies.[352] While a kinetic attack on Estonian infrastructure would uncontroversially constitute an armed attack for Article 5 purposes, NATO ultimately disagreed with Estonia's assessment that severity of the incoming cyberattacks could rise to a similar level of aggression; NATO's position with regards to these attacks offers tacit evidence that even these less discriminate cyberattacks are nonetheless observed to be less severe than kinetic options.

The case of Stuxnet offers further curious discrepancies in state responses to cyber aggression. Despite the destruction of its centrifuges and the setback of its nuclear aspirations, Iran's government offered comparatively little public outrage towards the violation of its sovereignty.[353] This is not to say that an Iranian response was wholly absent. In the wake of the attack, the Iranian government heavily invested in its cyber capabilities and has since been suspected as being behind a series of cyberattacks against Saudi and Qatari energy companies, as well as US financial institutions, in the early 2010s.[354] Accordingly, despite the physical nature of the harms inflicted by the Stuxnet worm, the Iranian response was limited to operations within cyberspace. By contrast, the Iranian government proved much more heated in its rhetoric

[351] Brian Orend, "Fog in the Fifth Dimension: The Ethics of Cyber-War," in *The Ethics of Information Warfare*, eds. Luciano Floridi and Mariarosaria Taddeo, vol. 14, Law, Governance and Technology Series (Cham: Springer International Publishing, 2014), 3–23, https://doi.org/10.1007/978-3-319-04135-3_1. 14-15.
[352] George R. Lucas Jr., "Emerging Norms for Cyber Warfare," In *Binary Bullets: The Ethics of Cyberwarfare*, eds. Fritz Allhoff, Adam Henschke, & Bradley Jay Strawser (New York, NY: Oxford University Press, 2016), 26.
[353] Randall R. Dipert, "Distinctive Ethical Issues of Cyberwarfare," In *Binary Bullets: The Ethics of Cyberwarfare*, eds. Fritz Allhoff, Adam Henschke, & Bradley Jay Strawser (New York, NY: Oxford University Press, 2016), 64.
[354] Ashish Kumar Sen, "Iran's Growing Cyber Capabilities in a Post-Stuxnet Era," *Atlantic Council*. April 10, 2015. Accessed March 31, 2022. https://www.atlanticcouncil.org/blogs/new-atlanticist/iran-s-growing-cyber-capabilities-in-a-post-stuxnet-era/.

following the drone assassination of one of its military generals by the US in early 2020; as news of the attack spread, the Iranian government decried it as an "act of international terrorism", promising it would take revenge against the US in response.[355] This rhetoric was supplemented with a kinetic response in the form of over a dozen Iranian ballistic missiles targeting a pair of Iraqi bases housing US troops.[356] Although the strikes ultimately failed to cause fatalities amongst US soldiers stationed at the bases, the kinetic attack nonetheless resulted in bodily harms, such as traumatic brain injuries, amongst base personnel.[357] While there are unignorable differences between the Stuxnet operation and the US drone assassination, the significant rhetorical discrepancy in the Iranian response and the state's more liberal deployment of kinetic force as a response in the latter case suggest kinetic operations are at least tacitly acknowledged to be more immediately severe than cyber alternatives. At the very least, the recognizable harms inflicted by conventional action seem to straightforwardly motivate kinetic responses in kind in a way that has hitherto not been seen with cyber operations.

Insofar as cyber operations pose less inherent risk of harm and the international community has expressed a general reluctance to declare a purely cyber operation as equivalent to an armed attack, I argue that the principle of last resort applies more loosely to cyber operations. While certain kinds of cyber operations may prove to be of equivalent severity to kinetic uses of force, these destructive cyberattacks are comparatively rare. Likewise, while some cyberattacks prove merely temporarily disruptive, there exists a wider range of attacks which appear more severe than economic sanctions. Accordingly, rather than serving as a functional equivalent in severity to either economic sanctions or kinetic uses of force, I suggest that cyber operations should instead be incorporated into the hiearchy of responses as an intermediary step, bridging less severe economic responses and more severe kinetic ones. While the principle of last resort necessarily restricts kinetic uses of force, the broader spectrum of

---

[355] Julian Borger and Martin Chulov, "US kills Iran General Qassem Suleimani in strike ordered by Trump," *The Guardian*, January 3, 2020. Accessed March 22, 2022. https://www.theguardian.com/world/2020/jan/03/baghdad-airport-iraq-attack-deaths-iran-us-tensions.

[356] Julian Borger and Patrick Wintour, "Missiles launched by Iran against US airbases in Iraq," *The Guardian*, January 8, 2020. Accessed March 31, 2022. https://www.theguardian.com/world/2020/jan/07/trump-iran-suleimani-threats-retaliation.

[357] Dan Lamothe, "These U.S. troops survived one of the greatest crises of the Trump era. A year later, they're still coping." *The Washington Post*, January 10, 2021. Accessed March 31, 2022. https://www.washingtonpost.com/national-security/us-military-iran-missile-attack/2021/01/10/651c3930-4fb0-11eb-b2e8-3339e73d9da2_story.html.

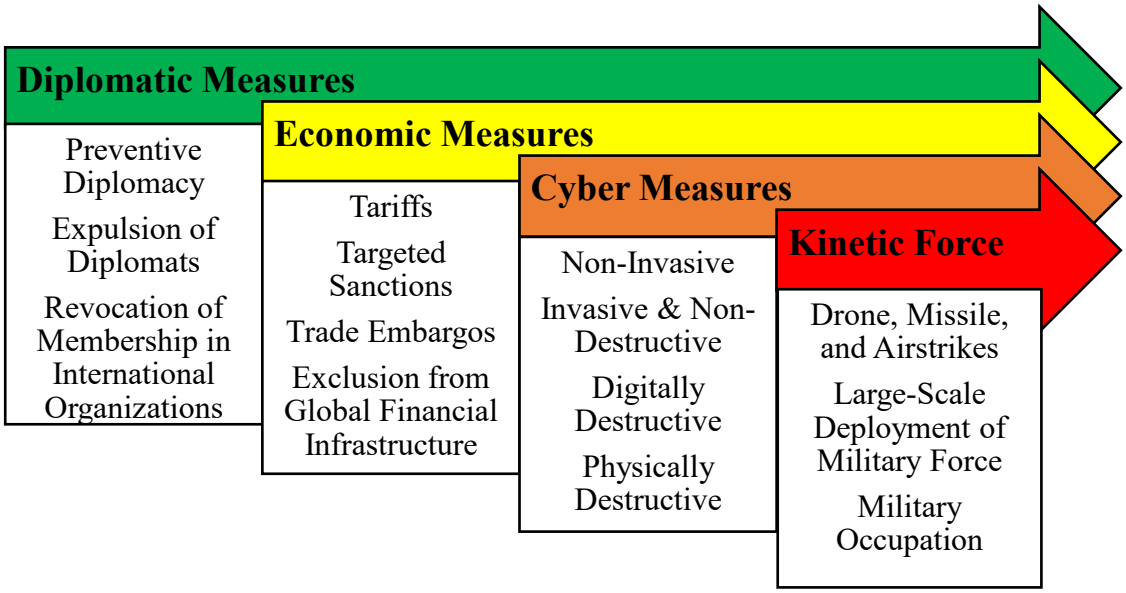
severity of cyberattacks allows for digital operations to be situated as measures of "later resort" rather than necessarily adhering to the same standards as kinetic alternatives.

To illustrate, suppose State A is building up its naval capabilities to project force into a region of territorially contested islands situated between it and State B. Following the unprovoked sinking of one of its patrol ships by State A's navy, State B begins to explore potential responses to curb further aggression by State A. Once diplomatic and economic efforts have proven unsuccessful, State B may decide to incapacitate State A's maritime navigation systems in order to drastically diminish State A's naval effectiveness and force it to negotiate. The decision to directly attack State A's maritime systems undoubtedly represents an escalation in response severity, however the degree of severity will vary depending on *how* State B seeks to achieve its objective. On the one hand, State B may launch drone strikes against a coastal base housing the majority of State A's naval navigation capabilities; doing so would successfully incapacitate the targeted systems, albeit at the cost of wider infrastructure damage to State A's naval base and potential injuries to its naval personnel. Alternatively, State B may deploy a logic bomb designed to wipe State A's navigation data and lock State A out of its own computer systems. While this alternative approach would deal a similarly large blow to State A's naval capabilities, it represents a visibly less severe response than the kinetic option. The discriminate nature of the attack ensures no adjacent infrastructure is physically destroyed, resulting in comparatively less cost imposed on State A. Likewise, the methodology of the attack avoids risking any needless injuries or deaths amongst State A's military personnel. The attack itself may be transient, with the erased data having been backed up by State B and restored in the event of an amicable settlement being reached between the belligerents. While the cyber approach is more severe than the preceding economic efforts, it nonetheless falls short of the threshold set by a kinetic attack.

This is not to say that cyber operations are inherently more or less severe than economic and kinetic responses. Rather, it would prove helpful to consider cyber operations as accommodating a broad spectrum of severity slotting between economic and kinetic options, with some overlap at either end of the scale. At the lower end, we may regard non-invasive, or invasive but non-destructive, cyber operations as being of roughly equivalent severity to some kinds of economic sanctions. While targeted sanctions may be less outwardly severe insofar as they target particular industries or individuals, broader economic sanctions which hit large

swathes of a state's economy could prove to be of similar severity to indiscriminate DDoS attacks which temporarily paralyze elements of a state's digital infrastructure. At the other end of the spectrum, cyberattacks resulting in kinetic-analogous harms may partially overlap with certain kinds of kinetic force. While a cyberattack achieving destructive results with a high degree of discrimination would be less severe than a kinetic attack towards the same ends, it is easy to imagine a cyber operation achieving destructive results through less discriminatory means, such as one which destroys maritime navigation systems by remotely detonating smart missiles located within the same naval base. This kind of attack could be treated as uncontroversially similar in severity to a straightforward kinetic alternative. Despite fringe cases of overlap existing, the perceived severity of most cyber operations appears to fall somewhere between economic sanctions and kinetic uses of force.

Figure 2. Cyber-Integrated Response Hierarchy



While the harms inherent to kinetic uses of force require them to be treated as measures of last resort, the comparative flexibility of cyberweapons allows for a loosening of the last resort principle for cyber operations. Despite the potential for cyberattacks to rise to the level of kinetic operations, current practices amongst states show that most cyberattacks fail to manifest in a physical capacity. Likewise, state rhetoric regarding past cyberattacks suggests cyberattacks have hitherto been regarded by the international community—almost as a matter of customary practice, we might say— as less immediately severe than kinetic alternatives, reducing the risk of

potential escalations of hostilities. Similarly, the wide range of potential objectives achievable by cyber operations and the varied means by which such objectives may be achieved renders it possible for states to tailor cyber operations to approximate the severity of both high-level economic sanctions and low-level kinetic operations. Given that the harms posed by most cyber operations fall short of those inflicted by kinetic operations, I argue that the threshold of last resort is lower for the former than it is for the latter. The lower last resort threshold for cyber operations allows for their deployment as a measure of *later* resort, following the exhaustion of diplomatic and economic options, but prior to the last resort deployment of more destructive kinetic force.

6.6 The Probability of Success Principle

The broader *jus ad bellum* framework of JWT is ultimately rounded out by the probability of success principle. In light of the harms inherent to the conduct of armed conflict, each of the preceding principles has worked towards limiting the situations within which states are justified in going to war, constraining the moral justifiability of war to solely the most exceptional of cases. To this end, the *jus ad bellum* framework has hitherto sought to ensure that any decision to go to war is 1) morally justified, 2) motivated by just intentions, 3) proportionate to the harm incurred, 4) publicly declared by the appropriate authorities, and 5) conducted only as a measure of last resort once less harmful (albeit still plausible) alternatives have been exhausted. Collectively, these principles introduce strict constraints on when war may be declared, mitigating the possibility of avoidable harm. Despite the general comprehensiveness of these principles, there nonetheless remains a critical consideration which remains unaddressed.

The probability of success principle adds to the *jus ad bellum* calculus by introducing the further demand that, should all other criteria be met, a belligerent only go war if it has reasonable prospects of success.[358] Insofar as war is inherently destructive, any state declaring war is aware that such a decision will necessarily require it to inflict harms against a belligerent, and, in turn, that it will likely incur harms of its own in the process. While unavoidable, these harms are typically viewed as the "cost of doing business", a necessary sacrifice made in pursuit of

[358] Orend, *War and Political Theory*, 91.

achieving a valued objective. As such, the resort to war by a state knowing, in advance, that it will be inevitably unsuccessful in achieving its objectives seems to accomplish little but the perpetuation of further, wholly avoidable, harm. In restricting moral justification for war to only those cases within which states have reasonable chances of success, this principle seeks to diminish the risk of needless conflicts and prevent the unnecessary harms that they entail.

This principle does not preclude the ability of smaller states to defend themselves against more powerful aggressors. While larger states are unlikely to suffer defeat in protracted armed conflicts with weaker belligerents, a "successful" outcome of conflict need not necessarily be constrained to the total annihilation of enemy forces. Wars may be justifiably fought in pursuit of smaller-scale, more attainable objectives. A small state is justified in standing its ground against a militarily superior aggressor in hopes of imposing costs sufficient for forcing the aggressor to either cease its offensive or become amicable to conditional negotiations. The Vietnam War saw Viet Cong forces resist against the much larger US military until mounting military losses and the domestic unpopularity of the conflict eventually motivated the withdrawal of US forces in 1973.[359] A similar phenomenon may be emerging once again with Russia's 2022 invasion of Ukraine, with Ukrainian forces inflicting unprecedented losses against the much larger Russian military, contributing to growing discontent amongst Russian citizens and officials.[360] While Ukraine may be unlikely to defeat the full strength of the Russian military, "success" in this case is centered upon the more realistic objectives of preventing the Russian military from achieving its own objectives and preserving Ukrainian sovereignty in the face of an existential threat. Insofar as a state's motivating objectives for war are realistically attainable, this principle can be fulfilled.

Although the principle of probability of success applies most directly to state decisions to go to war, it is also applicable to isolated usages of force which may not necessarily escalate to full-blown conflict. In keeping with the intention of mitigating unnecessary harm, individual uses of force are themselves only justifiable if, in addition to fulfilling the preceding principles, they enjoy reasonable chances of successfully achieving their objectives. A sizeable raid against an

[359] George C. Herring, "America and Vietnam: The Unending War," *Foreign Affairs* 70, no. 5 (1991): 104-119, https://doi.org/10.2307/20045006. 111.
[360] Anton Troianovski and Michael Schwirtz, "As Russia Stalls in Ukraine, Dissent Brews Over Putin's Leadership," *The New York Times*, March 22, 2022. Accessed March 23, 2022. https://www.nytimes.com/2022/03/22/world/europe/putin-russia-military-planning.html.

aggressor's military outpost is likely to cause physical destruction and bodily harm, however it enjoys reasonable prospects of successfully diminishing an aggressor's ability to project force within the region. By contrast, a small-scale raid launched with the intent of capturing a fortified military base is likely to violate the principle insofar as a small raiding party is exceedingly unlikely to succeed at assailing a fortified position housing an overwhelming number of armed personnel; in this latter case, the raid is unlikely to achieve anything beyond inflicting needless harms on both members of the raiding party and the defenders. It is worth addressing the possibility that the objective of a use of force is simply to cause harm to an inevitably victorious aggressor as sort of a "black eye" for its aggression; in such cases, simply causing an unignorable amount of harm may be seen as an operational success. While the probability of success principle might realistically be met in these cases, it is worth remembering that this principle is merely one of six meant to be taken in conjunction. Although these kinds of operation may hypothetically fulfill this principle, they would likely face greater difficulty fulfilling the rest, most notably that of right intention.

The principle of probability of success offers a final restriction within the *jus ad bellum* framework of JWT. By constraining permissible wars and engagements to solely those which enjoy reasonable chances of success, this principle works to mitigate unnecessary harms caused by conflicts that are likely to achieve little more than "death and destruction that will make no difference to the outcome".[361] While this principle initially appears to benefit powerful belligerents at the expense of weaker combatants, the definition of "success" with regards to this principle varies drastically from case to case; weaker states may target readily attainable goals falling well short of total victory, allowing them to justifiably deploy force towards these ends. Taken in conjunction with the remaining *jus ad bellum* principles of JWT, the probability of success principle seeks to ensure that states are only permitted to inflict harm in pursuit of realistically achievable ends, rather than granting moral backing to military actions appealing to morally desirable, but wholly unrealistic, goals.

[361] Orend, *War and Political Theory,* 91.

6.7 Lesser Risk, Same Reward: Revising Thresholds

Prior to any discussions regarding the probability of success principle's application to cyberspace operations, it is helpful to first divide the idea of cyberwar into two specific kinds of conflict based on the kinds of consequence each entails. The first of these is "pure" cyberwar. A pure cyberwar consists of a conflict within which combatants wage war using digital means towards predominantly digital ends. These conflicts may see belligerents deploy varying cyberweapons in pursuit of advancing their foreign policy objectives, however their effects will rarely manifest in any physical capacity; should the effects of a pure cyberwar cross the physical-digital divide, the consequences will typically be transient and non-destructive, akin to the effects of operations such as DDoS attacks. The second kind of cyber conflict is "hybrid" cyberwar. Unlike pure cyberwar, hybrid cyber conflict results in both cyber-specific and kinetic-analogous harms. Cyberattacks resulting in the scale and effects of kinetic attacks would fall under the purview of hybrid cyberwar. Insofar as the cyberattacks of hybrid cyberwar bridge the physical-digital divide in their effects, it is more likely that hybrid cyberwar would see kinetic responses to digital attacks; provided the harms inflicted by a cyber operation are suitably severe, kinetic measures may be deemed proportionate responses. Given that the expected consequences of these two kinds of cyber conflict differ drastically, so too should the threshold for probability of success vary in each case.

The less inherently harmful nature of cyberwar strategies allows for a less restrictive probability of success threshold in the case of pure cyberwar. Arguing in favour of this position, Jenkins notes that the principles of *jus ad bellum* ultimately gain their force from the severity of the harm they seek to prevent; insofar as war entails harm on an immense scale, the thresholds for these criteria must necessarily be set high to mitigate harm in all but the most serious, unavoidable cases.[362] Given that even the most discriminate of kinetic weapons entail severe harms, kinetic operations must necessarily have a high probability of success in order to be justifiably deployed. The deployment of kinetic measures in pursuit of unlikely objectives is likely to inflict further severe and lasting harms, with little corresponding likelihood of achieving meaningful progress towards the restoration of peace. In extreme cases, the needless deployment

[362] Ryan Jenkins, "Cyberwarfare as Ideal War," in *Binary Bullets: The Ethics of Cyberwarfare*, eds. Fritz Allhoff, Adam Henschke, & Bradley Jay Strawser (New York, NY: Oxford University Press, 2016), 108.

of kinetic weapons may motivate a dramatic escalation in force, ensuring even greater harm. Due to the severe costs associated with failure, kinetic weapons are only permissibly deployable when they are highly likely to achieve their objectives.

While the inherently ruinous nature of kinetic war limits the ability of kinetic operations to fulfill the probability of success criterion, there is more optimism regarding the potential of cyberweapons. To this effect, Jenkins asserts that "[the] less likely a war is to inflict indiscriminate harms, the lower the moral burden of proof to demonstrate probability of success".[363] Unlike kinetic weapons, cyber operations can be "maximally discriminate and perfectly reversible".[364] Cyberweapons can be designed to inflict severe harm on targeted systems, while simultaneously avoiding inflicting any adverse effects onto non-objective targets, thereby reducing the risk of serious collateral harm. Likewise, in the event of miscalculations in cyberwar, such as a worm infecting and damaging systems beyond those of its target or a cyberattack deployed against a state mistaken for the aggressor, the same cyberattacks may be swiftly recalled and their effects immediately reversed. While both kinetic and cyber wars are readily capable of achieving similar degrees of harm, cyberattacks have hitherto generally emerged as less harmful alternatives to kinetic operations. If the high thresholds of the *jus ad bellum* criteria are motivated by the severity of harm they are designed to ward off, then the lesser degree of harm inherent to cyber operations permits a loosening of the probability of success principle. Given the harms of pure cyberwar are generally less severe, pure cyberwar operations may be justifiably deployed with a lower probability of success than we would necessarily demand of kinetic alternatives.

A similar logic may be applied towards probability of success deliberations pertaining to hybrid cyberwar. The lesser degree of harm inherent to cyber operations renders them more permissible alternatives to conventional action. Should a non-destructive cyber operation and a discriminate kinetic attack enjoy similar probabilities of success, the cyber alternative proves more straightforwardly justifiable due to the diminished likelihood of serious harms resulting from failure to achieve its objective. In the event that the cyber alternative is itself destructive, it may nonetheless be given a lower probability of success threshold to clear; while a cyber operation may be incapable of removing the risk of collateral damage entirely, its more

[363] Jenkins, "Cyberwarfare as Ideal War," 108.
[364] Jenkins, "Cyberwarfare as Ideal War," 108.

discriminate methodology is nonetheless likely to result in fewer harms than would be inflicted by a kinetic alternative. As the gap in predicted degrees of harm between kinetic and cyber options narrows, so too would the probability of success threshold for cyber operations rise. However, provided there remains an appreciable gap between the predicted harms between kinetic and cyber alternatives, the novel methodology and effects of cyber operations allow for a lower probability of success threshold for cyber operations.

Setting a lower bar for probability of success in the case of cyberattacks complicates the potential interchangeability of cyber- and kinetic attacks within cyberwar. Insofar as hybrid cyberwar entails kinetic-analogous harms being wrought by cyber means, it is more likely that such conflicts would see kinetic measures deployed in response to cyberattacks, and vice versa. On the one hand, cyberattacks may prove perfectly reasonable responses in accordance with the probability of success criterion. Aggressive air operations may realistically be neutered by cyberattacks against an aggressor's air traffic control systems or hijacking an aggressor's own anti-air assets. Similarly, cyberattacks could bring an aggressive land operation to a halt by targeting vulnerable communication and logistics systems or disabling unmanned weapons systems. Accordingly, responses of a *different kind* remain justifiable insofar as they are demonstrably capable of achieving their objectives. Cyber responsese to cyber attacks are likewise justified by this principle as the former has appreciably realistic chances of succeeding in repelling the latter; a targeted state may disable the aggressor's cyber capabilities by targeting the offending computer systems or relevant network infrastructure with disruptive attacks, thereby halting further cyber aggression. Insofar as cyberweapons are capable of effectively responding to both kinetic and cyber uses of force, cyber strategies may fulfill the probability of success criterion in either kind of scenario.

Greater probability of success concerns revolve around whether kinetic operations may be justifiably deployed in response to cyber acts of aggression. The effectiveness of kinetic measures at curbing offensive cyber operations remains unclear. Kinetic responses to kinetic offensives meet the probability of success criterion insofar as they often target military infrastructure commonly known to be integral to fighting a kinetic war, such as airfields, military bases, and naval ports; airstrikes on an airfield hosting an aggressor's bomber aircraft enjoy a reasonable likelihood of preventing further bombing raids, thereby meeting the probability of success principle. However, cyber operations typically rely on an entirely different infrastructure.

Cyber operations do not require dedicated and identifiable military installations. Rather, they may readily be conducted from clandestine locations while exploiting "tight linkages between military and civilian networks", rendering it harder for responding states to discern between permissible and impermissible targets.[365] The disanalogous infrastructure underpinning cyber- and kinetic attacks renders it difficult to conceive of a kinetic operation which enjoys reasonable chances of directly stopping a cyberattack.

To illustrate, we may consider a case within which state-backed cyber operatives remotely seize control of a state's rail traffic networks and facilitate the derailment of a freight train transporting military equipment, causing damages of severe enough scale and effects for the targeted state to treat the operation as an armed attack. While effective cyber response options to this operation are readily discernible, it is much harder to formulate an effective kinetic response. A drone strike on a server farm would prove successful only if the aggressors were reliant on those servers; given the intertwined relationship between military and civilian computer networks, it is likely that the relevant infrastructure may not be targetable by such means. Perhaps a responding state may attempt to block internet access to an entire state by physically destroying network infrastructure across the aggressor's territory in a blanket effort to stop further aggression. This may likewise prove unsuccessful if the initial attack was conducted by a worm which has been preprogrammed to operate without further commands or triggers. Even if a hypothetical kinetic response with a realistic probability of success could be discovered, it is unlikely to have a notably higher chance of success than available by cyber means, or to do so without inflicting a significantly greater degree of harm.

As was the case with the principle of last resort, the probability of success criterion applies less strictly to cyber operations than it does with regards to kinetic alternatives. The costs of unsuccessful deployments of kinetic force are inescapably high given that kinetic action entails physical destruction and often permanent harms. In contrast, cyber operations are more discriminate and generally less harmful, except when specifically designed to mimic the kinds of effects inflicted by kinetic weapons. Insofar as the costs of an erroneous or unsuccessful cyber operation pale in comparison to a kinetic one, the former kind of force may be justifiably deployed with a lower likelihood of success than is required to motivate the latter. Furthermore,

[365] Edward T. Barrett, "Warfare in a New Domain: The Ethics of Military Cyber-Operations," *Journal of Military Ethics* 12, no. 1 (April 2013): 4–17, https://doi.org/10.1080/15027570.2013.782633. 10.

the probability of success criterion represents arguably the biggest hurdle for a state's ability to respond to cyber aggression with kinetic responses. In addition to the higher probability of success threshold needing to be met by kinetic operations, the idiosyncracies of cyber operations renders them particularly difficult to effectively respond to kinetically. The typical targets for kinetic operations, such as military bases, are absent; in their stead are tightly interwoven civilian and military networks and decentralized computer assets which may prove unrealistically targetable by kinetic means. This is not to say that such responses are never justifiable, however the situations within which they may permissibly deployed according to the probability of success principle are likely to be incredibly rare.

6.8 A Cyber *Jus ad Bellum*

Insofar as cyberattacks need not result in the same severe consequences which follow their kinetic counterparts, the consequentialist principles of *jus ad bellum* offer greater leniency to cyber operations. The principle of proportionality operates as a means of ensuring that initial aggression is not taken as an invitation for states to respond with disproportionate force. While both cyber- and kinetic attacks are capable of fulfilling the proportionality principle, the less harmful and more discriminate character of cyberattacks renders them typically more proportionate responses than kinetic alternatives. The principle of last resort adds further constraints by demanding that states exhaust non-forceful alternatives prior to resorting to force. Cyber operations are capable of a wider spectrum of harm than is accessible to kinetic attacks; while cyberattacks may feasibly rise to the level of armed attacks, the vast majority fail to be as severe. As states have hitherto proven reluctant to treat past cyber operations as severely as kinetic attacks, it would appear that cyber operations fall between economic sanctions and kinetic force as measures of later, but not last, resort. Finally, the probability of success principle seeks to prevent unnecessary harms brought about by pointless conflict. Provided that the strength of this principle is contingent on the severity of the harms it works to prevent, the less inherently harmful nature of cyber operations allows for a correspondingly lower probability of success. As the cost of getting it wrong, so to speak, is lower, these operations are more justifiably deployable than kinetic force.

Taken in conjunction with the preceding anti-consequentialist principles, these further criteria leave us with a comprehensive and satisfying cyber-specific conception of *jus ad bellum*. While tight-to-the-law approaches to cyberwar governance, such as the *Tallinn Manual*, are adept at evaluating cyber operations resulting in kinetic-analogous harms, their ability to accommodate the wider spectrum of potential cyber harms remains limited. This project turns towards JWT to build beyond existing LOAC evaluative frameworks to further address the kinds of cyberattacks which lay outside of the scope of efforts such as the *Tallinn Manual*. By analyzing each of the JWT *jus ad bellum* principles and identifying cyber-specific considerations for each criterion, this project has worked towards establishing the basis for a more comprehensive evaluative framework that not only accounts for kinetic-analogous cyberattacks, but also encompasses the wider range of cyber operations failing to manifest in physically destructive manners. Within the next chapter, we will consider past and hypothetical cyberattacks through both the LOAC and cyber-specific *jus ad bellum* frameworks to illustrate how the latter approach may serve as a stronger basis upon which to motivate interstate dialogue and develop norms of best practice with regards to cyberwar governance.

Chapter 7

Bridging the Governance Gap

7.1 Applying the Cyber *Jus ad Bellum* Framework

Having evaluated each of the six *jus ad bellum* criteria and identified cyber-specific considerations for each principle, we are now equipped with a broader evaluative framework for cyberattacks. This framework is not only applicable to cyberattacks resulting in kinetic-analogous harms, it also encompasses a fuller spectrum of cyber operations including those non-physically destructive cyber operations which are likely to remain the predominant state threats within cyberspace. In order to illustrate how this alternative account addresses the gaps of tight-to-the-law frameworks, we will now consider three different kinds of cyberattacks which prove disanalogous to conventional operations, whether by virtue of the specific types of harms they cause or the methodology by which harms are inflicted. In each case, we will identify the shortcomings of the LOAC approach and the ramifications they have on the ability of states to defend themselves against cyber aggression. We will subsequently examine the same case with the more flexible normative framework to show how the cyber *jus ad bellum* addresses the gaps left by the purely legal approach; and still preserves the inherent right to self-defense within cyberspace.

7.2 Familiar Harms, Unfamiliar Means: The Stuxnet Case

In June of 2010, the discovery of the Stuxnet worm made it starkly apparent that the effects of cyberweapons need not remain confined to the cyber domain. Devised as a part of Operation Olympic Games, a joint US-Israeli effort to derail Iran's controversial nuclear enrichment program, the Stuxnet worm managed to damage a significant number of centrifuges housed within the Natanz uranium enrichment facility to the point of inoperability.[366] Although the facility was air-gapped in the interests of protecting it from potential cyber interference, the

[366] Jon R. Lindsay, "Stuxnet and the Limits of Cyber Warfare." *Security Studies* 22, no. 3 (July 1, 2013): 365–404. doi:10.1080/09636412.2013.816122. 366.

worm infiltrated the plant aboard the compromised USBs of Iranian scientists, thereafter gaining access to the facility's control systems by exploiting a series of "zero-day vulnerabilities".[367] Central to the worm's targets were the programmable logic controllers which controlled the speed of the centrifuges used for the enrichment process.[368] Having infiltrated these control systems, the worm permitted the normal functioning of the "legitimate controller code" until an appropriate moment to strike appeared, at which point it surreptitiously halted the original code while launching replacement code written by the attacker.[369] This allowed the worm to manipulate the speed of the centrifuges over a period of time, forcing them to operate outside of their usual parameters, and ultimately resulting in their rapid degradation.[370] This diminished the number of functional centrifuges within the Natanz facility, thereby dealing a blow to Iran's ability to produce enriched uranium, an ingredient necessary for producing nuclear weapons.

While the Stuxnet worm did not bring a permanent halt to Iran's uranium enrichment program, it served to announce the arrival of cyberweapons as genuine means of inflicting tangible harms, such as bodily injury and physical destruction.[371] The immediate aftermath of the attack drew great speculation regarding the future of such cyberweapons, with proponents lauding the operation for its precision, while others, such as Russia's ambassador to NATO, expressed concerns that future such endeavours ran the risk of triggering unforeseeable disasters, such as a repeat of the Chernobyl incident.[372] The deliberations of the *Tallinn Manual* Experts were not immune to the controversies surrounding the Stuxnet worm. On the one hand, the worm was unanimously regarded by the Experts as constituting a use of force. Stuxnet was designed, and deployed, with the express intent of sabotaging Iranian uranium enrichment centrifuges. It was ultimately successful in this task. Insofar as "[a]cts that injure or kill persons or physically damage or destroy objects are uses of force", so too would Stuxnet constitute a use of force.[373] Consequently, based on the Experts' interpretation, a Stuxnet-like cyber operation could feasibly

[367] James P. Farwell and Rafal Rohozinski, "Stuxnet and the Future of Cyber War," *Survival* 53, no. 1 (February 2011): 23–40, https://doi.org/10.1080/00396338.2011.555586. 24.
[368] Farwell and Rohozinski, "Stuxnet and the Future of Cyber War," 24.
[369] Ralph Langner, "Stuxnet: Dissecting a Cyberwarfare Weapon," in *IEEE Security & Privacy*, vol. 9, no. 3, pp. 49-51, May-June 2011, doi: 10.1109/MSP.2011.67. 50.
[370] Farwell and Rohozinski, "Stuxnet and the Future of Cyber War," 24.
[371] Lindsay, "Stuxnet and the Limits of Cyber Warfare." 366.
[372] Lindsay, "Stuxnet and the Limits of Cyber Warfare." 366.
[373] Michael N. Schmitt and NATO Cooperative Cyber Defence Centre of Excellence, eds. *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations*. Second edition. Cambridge, United Kingdom; New York, NY, USA: Cambridge University Press, 2017. 333.

motivate a multilateral deployment of force in accordance with the provisions outlined within UN Charter Article 42.

No such consensus was forthcoming as the discussions turned towards whether Stuxnet earned the more severe distinction of armed attack necessary for invoking the right to unilateral self-defense stipulated in Article 51. The Experts explicitly referred to the Stuxnet case as being divisive on account of the general legal uncertainties regarding "the precise point at which the effects of a cyber operation qualify that operation as an armed attack".[374] For some of the Experts, the damage inflicted unto the Iranian centrifuges by the worm satisfied the scale and effects criteria necessary for regarding the operation as an armed attack against Iranian interests.[375] Accordingly, for one contingent of Experts, the consequences of a Stuxnet-like attack could feasibly justify self-defensive action. This opinion would not be shared by the remaining Experts; while conceding that the worm undoubtedly constituted a use of force, the remaining Experts staunchly maintained that the damages caused by Stuxnet were not sufficiently severe for warranting the stronger distinction of armed attack. The only Iranian elements to incur damages were the specific centrifuges targeted by the attack. The surrounding infrastructure was not damaged as a result of the worm's manipulation. Likewise, there was no corresponding injury or loss of life stemming from the sabotage effort. In the absence of significant collateral damage directly attributable to the attack, the operation remains a mere use of force. As a result, for the latter cohort of Experts, a state targeted by a Stuxnet-like operation would remain unjustified in responding to harms incurred with force of its own. The consequences would need to be significantly more severe to justify a forceful response.

In the case of Stuxnet, and similar such future cyberattacks, the absence of explicit legal thresholds identifying the specific point at which a use of force rises to the level of an armed attack serves to hamstring the ability of states to respond to cyber aggression which, in turn, allows the cyber aggressor to operate with greater impunity. In the absence of clear thresholds, the just cause criterion is only uncontroversially met in the worst of cases, those within which the damages are readily replicable by conventional armed attacks, such as air- or artillery strikes. Barring these (exceedingly rare) clear-cut exceptions, targeted states remain limited to less direct short-of-force responses, be they diplomatic or economic, while pleading their case for a

[374] *Tallinn 2.0*, 341.
[375] *Tallinn 2.0*, 342.

multilateral response in an international forum such as the UN Security Council. While diplomatic and economic sanctions may prove effective, their effects are not immediate; an aggressor may continue its operation for a prolonged period prior to incurring sufficiently deterrent costs. Likewise, the victim state may ultimately succeed in motivating a multilateral response, however it is unlikely that such an initiative would be quickly forthcoming. The general controversy surrounding severity evaluations of cyberattacks within this legal framework renders it unlikely that a consensus necessary for motivating a multilateral response is achievable in the event of cyberattacks less severe than a cyber-Pearl Harbor, let alone achievable within a tight timeframe following the initial cyberattack. Insofar as the Experts (operating within a familiar Western legal context) were incapable of reaching consensus regarding the severity of Stuxnet, it is unlikely that a consensus would be easily achieved between diametrically opposed member states with veto powers.

While the uncertainties present within the tight-to-the-law evaluative approach render such approaches overly restrictive on the ability of states to respond to cyberattacks resulting in discriminate destruction, the flexibility of a cyber-specific framework offers a streamlined avenue towards evaluating, and responding to, the original act of cyber aggression. Under the provisions outlined for a cyber *jus ad bellum*, the Stuxnet worm would satisfy the just cause criterion for Iran. As Smith notes, insofar as Iran took concerted measures to protect their enrichment facilities from foreign intrusions (including designing the entire facility to be air-gapped), the Stuxnet worm could be regarded as an armed attack against Iran's core interests.[376] The Stuxnet worm was not an effort to shape Iranian deliberations towards a specific political outcome, nor was it an imposition of cost meant to dissuade the state from pursuing its nuclear enrichment program. Rather, the worm was devised by US & Israeli intelligence specifically as a means to "[bypass] Iranian deliberations and judgments to impose a particular outcome upon them".[377] Accordingly, it is irrelevant whether the Stuxnet operation proved ultimately less destructive than a conventional use of force deployed towards the same objective. Insofar as it

[376] Patrick Taylor Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account: Cyberattacks as Casus Belli," *Journal of Applied Philosophy* 35, no. 2 (May 2018): 222–41, https://doi.org/10.1111/japp.12169. 234.
[377] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account: Cyberattacks as Casus Belli," 234.

represents a direct imposition of will designed to bring about a specific political objective, it satisfies the just cause criterion.[378]

Although the just cause criterion is more readily satisfied within the cyber *jus ad bellum* framework than within the tight-to-the-law approach, this does not mean that potential Iranian responses to Stuxnet would be unconstrained. Any forceful Iranian response to Stuxnet would need to fulfill the principle of last resort, with the Iranian government exhausting diplomatic and economic alternatives, or otherwise determining them to be ineffective at curtailing further aggressive interference. Given the economic disparity between the US and Iran, it is entirely likely that the latter may be justified in forgoing such softer measures entirely in favour of more forceful countermeasures. In the interests of preventing undue escalations of hostilities stemming from a weaker threshold for just cause, the remaining *jus ad bellum* principles work to constrain the kind of force potentially deployable as a response to a Stuxnet-like operation.

Firstly, the kinds of permissibly forceful operations at Iran's disposal are inherently limited by the principle of proportionality. While the *jus ad bellum* framework may morally permit a forceful response, the force of the response cannot itself be disproportionate to the harms incurred. While the Stuxnet worm was effective at hindering Iran's nuclear enrichment program, its effects were not permanent, nor did recovery demand an extensive period of rebuilding; Iran's enrichment program was largely restored over the course of the next year.[379] While the degradation of the centrifuges ultimately caused physical damages, the damages inflicted did not include extensive damage to the physical infrastructure of the Natanz facility, nor did it include personal injury to the facility's operators. The precision of the damages caused by Stuxnet, and the lack of corresponding collateral harms, renders the operation dramatically dissimilar from conventional deployments of force, such as air- and drone strikes, which regularly inflict greater degrees of harm.

Consequently, despite the fulfillment of the just cause criterion, the proportionality principle would prohibit the deployment of Iranian kinetic force as a response to the worm. Although Stuxnet resulted in physical damages which are in theory replicable by conventional force, it is difficult to conceive of a kinetic operation that offers a realistic means of inflicting

[378] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account: Cyberattacks as Casus Belli," 234.
[379] Lindsay, "Stuxnet and the Limits of Cyber Warfare." 366.

proportionate harms in this case. Rocket attacks are highly indiscriminate and likely to inflict disproportionate harms in the form of widespread physical destruction. Drone strikes, while more precise, nonetheless result in tangible harms to infrastructure and personnel in the vicinity of the target. A covert special forces raid may present a hypothetically proportionate response if, during the raid, only the specific target is destroyed without the infliction of collateral damage; however, such operations carry a high likelihood of personal injury and/or death if the target of the raid is itself guarded by armed personnel. As a result, potential kinetic responses to Stuxnet inherently raise concerns of proportionality insofar as the discriminate nature of the worm is unparalleled amongst realistic conventional operations. A use of proportionate force is far more likely to emerge in the form of a cyber operation than it is a conventional response. For example, a cyberattack derailing a US weapons development program by sabotaging key testing equipment would qualify as a proportionate use of force insofar as it would inflict a similar degree of damage, without being likely to motivate any corresponding escalatory harms. While we cannot fully discount the possibility of a powerful state responding with escalatory force motivated by hurt pride, such responses would be flagrant violations of the cyber *jus ad bellum* and potentially leave the responding state vulnerable to international sanctions or other punitive measures. Insofar as the physical harms caused by a Stuxnet-like attack are highly discriminate, only cyberattacks represent realistically proportionate responses.

Secondly, the reasonable probability of success criterion further restricts Iran's ability to respond to Stuxnet-like weapons with conventional force. Let us assume that any forceful response by Iran in this case is motivated by the correct intention, namely that of deterring further aggression. For any such response to satisfy the probability of success principle, it must enjoy a reasonable likelihood of succeeding at preventing further foreign aggression directed against its core national interests. In theory, this may be achieved either by causing similar damages to the aggressor to impose a cost upon them that may disincentivize further acts of aggression or by specifically targeting elements of the aggressor's cyber capabilities to prevent them from being able to launch similar attacks. In either case, it is difficult to propose a conventional response which would succeed at either imposing a sufficient cost on an aggressor or meaningfully hinder its ability to conduct future cyber operations while still satisfying the preceding principle of proportionality. A rocket barrage launched against an aggressor's power grid in response to a Stuxnet-like operation would be undeniably effective at ensuring would-be

aggressors think twice before deploying uses of force, however such a response would undoubtedly serve as a gross violation of the proportionality principle.

Beyond compounding concerns of proportionality, there are also major questions regarding the *effectiveness* of conventional responses to cyber offenses. The covert and decentralized infrastructure underpinning cyber operations renders them particularly difficult to address with any conventional means, let alone proportionate ones. A state's ability to launch cross-border drone operations can be removed by destroying airfields and hangars housing the drones. By contrast, cyber operations rarely have comparable designated military or intelligence infrastructure and, in many cases, also make significant use of existing civilian internet infrastructure—ranging from electricity grids, to satellite links, to internet service providers, down to the raw fiber optic wiring. As a result, there is a noted lack of hard targets that, if struck, would result in a significant blow to a state's (but *only* the state's) cyber capabilities. Accordingly, any Iranian deployment of kinetic force in response to Stuxnet remains highly unlikely to satisfy the demands of reasonable probability of success. By contrast, we may imagine an Iranian cyber response in the form of an encryption attack against US intelligence agencies, locking US cyber operators out of their systems while deleting partial code intended for use in future cyberweapons. This form of response would have a reasonable likelihood of success in dealing a blow to the US' cyber capabilities, while nonetheless adhering to the principle of proportionality due to its discriminate nature.

The Stuxnet worm poses a problem for the LOAC evaluative approach for cyberattacks insofar as the discriminate nature of its methodology renders it difficult to draw severity comparisons between it and potential kinetic analogues. As a result, the tight-to-the-law approach suggests that a Stuxnet-like operation fails to fulfill the just cause criterion, restricting the targeted state's ability to respond to incurred cyber force unilaterally. This serves, perhaps perversely, to increase the effectiveness of highly discriminate and physically destructive offensive cyber operations. Likewise, this also emboldens aggressors, who may operate more aggressively with the knowledge that all but the most severe of cyberattacks will fail to morally justify potential retaliations. In contrast, the cyber *jus ad bellum* approach readily recognizes Stuxnet-equivalent attacks as constituting just cause for a unilateral, potentially forceful, response in the interests of self-defense. While the satisfied just cause criterion may morally justify the deployment of defensive force in response to Stuxnet, the nature of this force is

heavily qualified by the further principles of proportionality and probability of success. Consequently, while a cyber response is morally justifiable in this case, kinetic retaliation remains morally impermissible. As a result, the *jus ad bellum* approach to Stuxnet preserves the inherent right of the state to defend itself in the event of aggression (regardless of whether it is conventional or cyber in nature), while continuing to mitigate the risk of escalation triggered by disproportionate responses.

7.3 Disinformation, Disillusion, Division in Ukraine

On November 21, 2013, Ukraine's hopes of gaining European Union (EU) membership were dashed as then-President Viktor Yanukovich opted to renege on the Ukraine-EU Association agreement due to be signed at the EU Eastern Partnership Summit in Lithuania a week later.[380] Yanukovich's decision triggered widespread domestic unrest, with thousands of pro-EU demonstrators taking to the streets in what would become known as the Euromaidan protests. On February 20, 2014, following the threat of withheld Russian aid should the unrest be allowed to persist, Yanukovich's government deployed lethal force against the protestors.[381] Facing the threat of an armed coup following his government's failed efforts at quelling the demonstrations, Yanukovich fled to Russia on the 21st of February and an interim Ukrainian government was formed to restore stability to the nation.[382] As its grip on Ukraine's central government weakened, Russia would turn its attention towards strengthening its hold on the strategically relevant Crimean Peninsula.

Russia's subsequent annexation campaign of the peninsula displayed an explicitly hybrid character. On the 27th of February, "little green men" bearing Russian arms and military uniforms (albeit bereft of identifying military insignia) would seize key strategic points throughout the Crimean region.[383] Amongst these targets were integral elements of Crimea's telecommunications infrastructure, as armed personnel stormed and occupied Ukrainian

[380] Jan Stinissen, "A Legal Framework for Cyber Operations in Ukraine," in *Cyber War in Perspective: Russian Aggression against Ukraine*, ed. Kenneth Geers, NATO Cooperative Cyber Defence Centre of Excellence, 2015, https://www.ccdcoe.org/library/publications/cyber-war-in-perspective-russian-aggression-against-ukraine/. 125.
[381] Scott Jasper, *Russian Cyber Operations: Coding the Boundaries of Conflict*, (Washington, DC: Georgetown University Press, 2020). 55.
[382] Stinissen, "A Legal Framework for Cyber Operations in Ukraine," 125.
[383] Jasper, *Russian Cyber Operations: Coding the Boundaries of Conflict*, 56.

television and radio stations, while also "severing cables and routing calls through Russian mobile operators".[384] These kinetic efforts would strengthen Russia's influence within the digital domain, as the captured infrastructure was repurposed by Russian operatives to "control public opinion and indirectly shape decisions in favor of Russia".[385] To this end, this infrastructure was used as a conduit for Russian propaganda, with Ukrainian channels being replaced with Russian programming, while Crimeans had their access to outside sources of information restricted.[386] This allowed Russia to effectively shape the narrative surrounding the Euromaidan protests and the ousting of Yanukovich's government, strengthening its ability to portray the incoming Ukrainian government as "illegitimate" and posing an existential threat to ethnic Russians residing within Ukraine.[387] These efforts, working to exacerbate the existing rift between the Crimean Peninsula and the Ukrainian central government in Kyiv, ultimately enabled Russia to achieve "a near bloodless coup de main" as they annexed the region in 2014.[388]

The hybrid nature of the Russian operation within Crimea places it firmly within the evaluative capabilities of LOAC frameworks. In a retrospective analysis of the 2014 cyber operations in Ukraine, Stinissen, himself a member of NATO's Cooperative Cyber Defence Centre of Excellence, asserts that cyber operations will typically be deployed as "a facilitator for other, more traditional types of warfare", rather than standalone operations.[389] Accordingly, for Stinissen, the appropriate legal framework for a cyber operation is contingent on the context of the broader conflict; insofar as cyber operations are deployed towards waging a particular conflict, the laws "applicable to the conflict as a whole should be applied to the cyber activities that are part of it".[390] Notably, this approach requires that a given conflict first rises to the level of an armed conflict so that the LOAC become applicable. With regards to the cyber operations conducted within Ukraine from late 2013 and onwards, Stinissen's analysis of the unrest is

---

[384] Glib Pakharenko, "Cyber Operations at Maidan: A First-Hand Account," in in *Cyber War in Perspective: Russian Aggression against Ukraine*, ed. Kenneth Geers, NATO Cooperative Cyber Defence Centre of Excellence, 2015, https://www.ccdcoe.org/library/publications/cyber-war-in-perspective-russian-aggression-against-ukraine/. 62.
[385] Marie Baezner, "Cyber and Information Warfare in the Ukrainian Conflict," *Center for Security Studies (CSS), ETH Zürich*, (October 2018), https://doi.org/10.3929/ETHZ-B-000321570. 14.
[386] Baezner, "Cyber and Information Warfare in the Ukrainian Conflict," 14.
[387] Benjamin Jensen, Brandon Valeriano, & Ryan Maness, "Fancy Bears and Digital Trolls: Cyber Strategy with a Russian Twist," *Journal of Strategic Studies* Vol. 42. Issue 2. (January 2019), accessed May 29, 2021. https://doi.org/10.1080/01402390.2018.1559152. 14-15.
[388] Jasper, *Russian Cyber Operations: Coding the Boundaries of Conflict*, 58.
[389] Stinissen, "A Legal Framework for Cyber Operations in Ukraine," 123.
[390] Stinissen, "A Legal Framework for Cyber Operations in Ukraine," 123.

broken into three specific phases. The first phase, that of the events of the Euromaidan protests, is ultimately regarded by Stinissen as an internal matter falling below the level of an armed conflict. While there were cyber operations conducted by both state and non-state entities within Ukraine, the conflict was nonetheless considered by Stinissen to be "primarily an internal matter between a state and an opposition within that state".[391] As a result, the cyber operations at this stage are considered a matter of domestic law, rather than falling under the purview of the LOAC.

Following the culmination of the Euromaidan protests and the ousting of Yanukovich, the Ukraine crisis would enter its second stage with the creation of the interim Ukrainian government and the eventual annexation of the Crimean Peninsula. The second stage saw suspected Russian forces deploy to secure strategic points throughout Crimea and prevent the intervention of the Ukrainian armed forces, while there was a concurrent uptick in cyber operations targeting both private and public Ukrainian digital assets.[392] For Stinissen, it is this second stage of the crisis which brings the broader conflict into the fold of the LOAC insofar as the unilateral deployment of armed forces into the territory of a sovereign state constitutes a clear use of force.[393] Once this threshold is met, the LOAC then "regulates the conduct of all actors in the conflict, including the cyber actors".[394] Accordingly, the usual rules governing the deployment of force, such as those prohibiting engagement with certain targets, apply equally to cyberattacks and kinetic strikes.

Within his analysis, Stinissen employs the *Tallinn Manual*'s definition of a forceful cyberattack, namely those cyber operations which are "reasonably expected to cause injury or death to persons or damage or destruction to objects".[395] The cyberattacks which took place within Ukraine in 2014 failed to manifest in these sorts of consequences; most cyber operations within the region could be classified as instances of either cyber espionage or information operations. In the first case, Stinissen notes that remotely conducted espionage, such as that conducted with cyber means, fails to violate the LOAC. With regards to information operations, Stinissen remarks that cyberattacks of this variety are "not directly addressed in the Law of

[391] Stinissen, "A Legal Framework for Cyber Operations in Ukraine," 125.
[392] Stinissen, "A Legal Framework for Cyber Operations in Ukraine," 126.
[393] Stinissen, "A Legal Framework for Cyber Operations in Ukraine," 127.
[394] Stinissen, "A Legal Framework for Cyber Operations in Ukraine," 131.
[395] Stinissen, "A Legal Framework for Cyber Operations in Ukraine," 132.

Armed Conflict" and, as such, are subject to a greater degree of interpretation to determine whether they may pose some sort of violation.[396] While acknowledging that a state-backed cyber effort to disrupt democratic elections may feasibly constitute a transgression in the eyes of international law, Stinissen nonetheless maintains that such an effort would fail to specifically violate any LOAC.[397] As a result, barring those rare attacks resulting in physical destruction, cyberattacks within conflict typically fall short of drawing real scrutiny from the established LOAC.

These interpretations of just cause hold interesting ramifications for the third stage of the Ukraine conflict, namely the outbreak of hostilities within eastern Ukraine following Russia's annexation of Crimea. In April of 2014, pro-Russian separatist groups took control of television and radio stations throughout the Donbas, alongside destroying communications and broadcast infrastructure so as to "isolate the region from Ukrainian media, communication, and financial services",[398] enabling a tight control over the information sphere within the region. At the same time, widespread cyber operations against the rest of Ukraine have continued, hindering the state's ability to respond effectively to growing unrest within the country's eastern region. While the situation within the Donbas has since been escalated to the level of an international armed conflict following the invasion by Russian armed forces in early 2022, it was, prior to that moment, determined to be a "noninternational armed conflict" between the Ukrainian state and separatist forces by the International Committee of the Red Cross, a position which is echoed by Stinissen within his analysis.[399] In the absence of Russia exercising "overall control" over the pro-Russian forces via "organising, coordinating, and planning their operations", the conflict was deemed a non-international armed conflict not directly governed by the LOAC.[400]

While the LOAC approach ultimately determined that Russia's efforts within Crimea (and, later, the Donbas) constituted a use of force, this distinction was largely contingent on the physical *boots-on-the-ground* presence of Russian military personnel. Despite the physical component of Russia's involvement having been preceded by an information war heavily reliant on cyber operations, the lack of kinetic-analogous harms wrought by these cyberattacks ensured

[396] Stinissen, "A Legal Framework for Cyber Operations in Ukraine," 133.
[397] Stinissen, "A Legal Framework for Cyber Operations in Ukraine," 133.
[398] Jasper, *Russian Cyber Operations: Coding the Boundaries of Conflict*, 62.
[399] Stinissen, "A Legal Framework for Cyber Operations in Ukraine," 130.
[400] Stinissen, "A Legal Framework for Cyber Operations in Ukraine," 130.

that they would fall well short of the use of force threshold necessary for the LOAC to become applicable.[401] Accordingly, the LOAC failed to directly govern the cyber operations conducted within the scope of the conflict prior to physical engagement because the conflict was: 1) deemed a domestic issue; and 2) the cross-border cyberattacks failed to qualify as uses of force due to the absence of physical harm. The information operations launched in support of the annexation are treated as less severe threats which, while working to strengthen pro-Russian sentiment and mistrust towards Kyiv within Crimea, are not themselves sufficient for constituting uses of force. This is further evidenced by Stinissen's analysis of the conflict in Donbas shortly after the annexation of Crimea, as, despite the cyber involvement of Russian military and intelligence organizations in support of pro-Russian separatists,[402] the conflict would remain classified as a noninternational armed conflict up until Russian soldiers were physically deployed into the region. As a result, while the LOAC would be effective at governing *physically destructive* cyber operations within Crimea and Donbas, they are considerably less effective as a tool for governing operations falling short of that exceedingly high threshold for use of force.

While the eventual deployment of Russian troops into Crimea and Donbas finally resulted in both operations being declared violations of international law, conceiving of international conflicts as necessarily requiring a physical use of force omits the possibility that non-destructive *pure* cyberwar could rise to the level of international conflict. Although Russia's digital efforts within Crimea were supplemented with physical seizures of relevant communications infrastructure, it is not beyond the realm of imagination to suggest that Russia could feasibly seize control of that region's information landscape—and perhaps beyond—using purely cyber means, particularly as society grows more reliant on digital infrastructure and offensive cyber capabilities continue to evolve. Rather than physically capturing strategic points, Russia may instead deploy more sophisticated cyber operations, such as cryptographic attacks on telecommunications infrastructure throughout Ukraine, preventing access by Ukrainian operators and allowing for unfettered Russian access to broadcast networks to facilitate the dissemination

[401] Stinissen, "A Legal Framework for Cyber Operations in Ukraine," 133.
[402] Jasper, *Russian Cyber Operations: Coding the Boundaries of Conflict*, 64.

of propaganda and control of the information sphere within politically volatile contested regions, such as Crimea or the Donbas.[403]

At the same time, disruptive cyber operations may be launched at the central government within Kyiv, restricting its ability to combat misinformation and communicate with citizens or government officials residing within the area. In volatile regions, these kinds of information operations can serve to dramatically undermine the common life holding various culturally distinct regions within a state together; in the most extreme of cases, such operations may contribute to the rise of irredentist or separatist factions within the targeted state, significantly destabilizing it. Within the Donbas, this sort of political misinformation was employed to magnify residents' feelings of abandonment by the central government in Kyiv following the events of the Euromaidan protests and subsequent regional violence; efforts to amplify these sentiments can serve to "stimulate a new sense of identity, especially one along regional lines, strengthening political alienation among Donbas citizens and possibly leading them to side with the separatists of the DNR [Donetsk People's Republic] and LNR [Luhansk People's Republic]".[404] In addition to softening local sentiments towards separatist movements amongst those without strong pro-Russian leanings, concerted information operations disseminating biased sources of information also serve to reinforce pro-Russian beliefs amongst those who already held such sensibilities,[405] further fracturing the region's sense of belonging within the greater Ukrainian state and potentially motivating greater extremism amongst separatists.

Under the LOAC approach, these efforts would fall short of uses of force. None of the cyberattacks employed result in any physical damage or bodily harm. Likewise, most of the immediate effects of the cyber operations may be readily reversible, restoring full Ukrainian control of its own network infrastructure. In the absence of a physical operation deploying military personnel into sovereign territory to capture communication infrastructure, the broader campaign appears to employ no force (as defined in tight-to-the-law frameworks) in pursuit of its

---

[403] Neil Rowe, *Towards Reversible Cyberattacks*, Reading: Academic Conferences International Limited, 2010. http://search.proquest.com.proxy.lib.uwaterloo.ca/conference-papers-proceedings/towards-reversible-cyberattacks/docview/869507120/se-2?accountid=14906. 263.

[404] Elise Giuliano, "Who Supported Separatism in Donbas? Ethnicity and Popular Opinion at the Start of the Ukraine Crisis," *Post-Soviet Affairs* 34, no. 2–3, May 4, 2018: 158–78, https://doi.org/10.1080/1060586X.2018.1447769. 175.

[405] Leonid Peisakhin and Arturas Rozenas, "Electoral Effects of Biased Media: Russian Television in Ukraine," *American Journal of Political Science*, 62(3), January 19, 2018: 535-550, http://dx.doi.org/10.2139/ssrn.2937366. 30.

objective. This does not hold true under the cyber *jus ad bellum* framework here developed. Despite the absence of physical intervention, this kind of concerted campaign of cyber-facilitated information warfare represents the clear efforts of a foreign state to achieve a specific political objective, whether the annexation of a strategic region or the fracturing of a rival state to undermine its domestic security. Insofar as Russia's cyber operations sought to achieve these objectives through direct manipulation of the Ukrainian common life (in particular, that of Ukrainians residing within Crimea and the Donbas) in hopes of dispelling a common Ukrainian identity and potentially escalating domestic tensions to the level of civil war, these kinds of cyber operations likewise satisfy the just cause criterion.

As with the Stuxnet case, the lower threshold for just cause is tempered with constraints on the type of forceful response that may be permissibly deployed in the form of the remaining principles. Assuming the state fulfills the principle of right intention insofar as it intends to prevent further Russian aggression, any Ukrainian response would need to fulfill the principle of last resort, first exhausting non-forceful measures, such as diplomatic and economic sanctions, prior to escalating to more forceful alternatives should these prove unsuccessful. Once non-forceful options have been exhausted or deemed likely ineffective, Ukraine may then explore responding to the purely cyber aggression with force of its own. While the principle of just cause alone may not differentiate between cyber and kinetic responses, it would be difficult to formulate a kinetic Ukrainian response which is both proportionate to the harm incurred, and that enjoys a reasonable probability of successfully achieving its objective. In the first case, the harms inflicted by kinetic countermeasures would be radically different from the non-physical harms directly inflicted by information warfare cyberattacks; the disanalogous nature of these harms renders proportionality comparisons between the two largely untenable and makes it more likely that a kinetic response will result in further kinetic retaliations than being seen by the initial aggressor as a proportionate response to their own aggression. For the second principle, a kinetic response to cyberattacks is unlikely to have the direct result of stopping the cyberattacks; the infrastructure required for an encryption attack is minimal and potentially impossible to target with conventional means. Despite this, the cyber *jus ad bellum* would permit Ukraine to respond to such digital uses of force with digital force of their own, allowing them to exercise stronger and (perhaps most importantly) more immediate responses than merely diplomatic or economic sanctions.

In the case of the Ukraine conflict, the reliance of the LOAC approach on the presence of a physical component bestows too strong of an advantage on the aggressor. In the absence of a boots-on-the-ground element, Russia's operations within Crimea and the Donbas would fall outside of the purview of the LOAC. Accordingly, insofar as such operations are deemed short of uses of force, Ukraine's ability to respond to Russia's destabilizing efforts is limited to non-forceful measures, such as diplomatic and economic sanctions, both of which have historically proven ineffective in the country's dealings with its neighbour. By contrast, the cyber *jus ad bellum* framework more readily identifies non-destructive, purely digital, destabilization operations as potentially constituting aggressive uses of force. This approach not only identifies these cyber operations by a foreign belligerent as impermissible actions constituting grievous violations of sovereignty, but it also loosens the restrictions placed upon the ability of cyber victims to respond to cyber aggression lacking a physical component. In the case of Ukraine, identifying Russia's efforts in the cyber domain as uses of force permits Ukraine to justifiably deploy forceful cyber countermeasures. This allows for a response that is more immediately effective than sanctions, potentially closing the window of effectiveness of the original cyberattack, thereby mitigating the extent of its damages. Additionally, recognizing the forceful nature of Russia's cyber efforts in Crimea and Donbas also serves a pragmatic function in categorizing these kinds of cyber operations as being generally impermissible, potentially contributing to a new norm of cyber conduct within which invasive cyberoperations of this kind are emphatically met with international denouncement, thereby limiting their future use.

7.4 Invasive Disruptions: A Hypothetical Case

Suppose a series of small coastal states find themselves facing severe climate change risk in the form of soil erosion and flooding projected to render significant swathes of current territory uninhabitable, forcing the relocation of thousands of climate migrants within the near future. Publicly pledging its support to these vulnerable neighbouring states, State A adopts a series of policies to both take in many of these refugees, as well as to facilitate the transportation of the remainder through its territory to like-minded bordering states that have expressed a similar desire to assist. To this end, State A invests heavily in its digital logistics infrastructure to bolster its ability to efficiently relocate an increased volume of migrants. At the same time, State

A grants contracts to domestic logistics firms to assist in both documenting and processing new migrants, as well as coordinating their relocation throughout State A and neighbouring states. These preparations are soon put to the test, as rising sea levels dramatically diminish hospitable land, displacing tens of thousands of climate refugees who begin filtering through State A's territory.

As State A begins its intake of climate migrants, State B voices its discontent with State A's efforts. Citing past attacks on its interests by extremist organizations originating within the climate-threatened countries, State B asserts that the mass climate migration being facilitated by State A poses an imminent threat to State B's national security insofar as extremist combatants may pose as refugees to position themselves throughout State A and other nations bordering State B. This, State B argues, would introduce major challenges to its ability to defend against the extremist threat and increases the likelihood of direct extremist violence against State B's citizens and territory. Motivated by these concerns, State B explores both diplomatic and economic measures designed to urge State A to reconsider its present climate migration policies. Ultimately, State B's initial response takes the form of a diplomatic withdrawal from State A, alongside targeted sanctions on trade between the two states.

Following the failure of these economic and diplomatic measures due to State A dwarfing the influence of State B on both fronts, the latter resorts to a large-scale cyber campaign against the former with the intention of undermining State A's ability to efficiently accommodate and transport the increased volume of migrants. To this end, State B launches penetrative cyberattacks against both state- and private infrastructure within State A's territory, causing mass disruptions throughout the country. Some of the resulting disruptions are centered on transportation infrastructure. State A's rail infrastructure grinds to a halt as rail control systems are taken offline, preventing train monitoring and communication between conductors and controllers. Airline scheduling systems and maintenance records are likewise targeted, grounding most flights within State A's airspace. At the same time, further disruptive attacks target local transit organization systems, as well as ride sharing applications, in order to significantly disrupt mobility throughout State A at every level. Collectively, these attacks are deleterious for State A's ability to efficiently transport migrants, functionally stranding them throughout the state's territory.

Further disruptions target databases storing information integral to State A's relocation efforts. The databases of logistics contractors affiliated by State A are compromised, resulting in targeted retractions of records and attestations of climate migrants. This undermines the efficiency of State A's migration plan on two fronts. Firstly, retracting access to these records may remove key data required for entry to other states, such as proof of education, financial and criminal records, and so on. While the original copies of these records may exist in the hands of the migrants, disruptions to these online repositories significantly set back the migration process, as neighbouring states may refuse entry until these records are restored as a security measure. This results in a much greater number of migrants needing accommodation within State A as the borders are plagued by longer processing times. Secondly, these attacks also target cyber assets used for facilitating accommodation of migrants within State A, ranging from government-run housing to private services such as short-term rental platforms. Disrupting access to these elements dramatically undermines State A's ability to provide shelter and general aid to the migrants within its territory, potentially leaving them homeless for a prolonged period as efforts are made to restore the regular functionality of these systems. In the interim, the larger-than-expected volume of migrants left stranded within State A is likely to strain certain state functions, such as healthcare and law enforcement, potentially causing further problems domestically.

Collectively, these cyber efforts serve as a clear imposition of will by State B onto State A. While State A had organically self-deliberated on its decision to adopt humanitarian policies regarding climate migration, State B's ongoing severe and invasive cyber operation ensures that State A cannot act in accordance with its own political will. Moreover, State B's efforts ensure that there will be deleterious effects not only for the migrants themselves, but also the residents of State A as the state incurs greater-than-expected costs and dislocations associated with the unforeseen stranding of migrants. It is easy to imagine that State B may further contribute to domestic unrest within State A by supplementing these disruptive efforts with concerted information campaigns on social media, attempting to shift public perception of the migrants within State A and turn public sentiment against them. Ultimately, the objective of these cyberattacks is to motivate a policy reversal by State A with regards to its climate migration initiatives, achieving what the diplomatic and economic efforts were incapable of doing.

While State B's cyber efforts serve to manufacture a crisis within State A's territory, it remains uncertain whether State A has the option of a forceful response at their disposal under the LOAC framework offered by the *Tallinn Manual*. On the one hand, these attacks are evidently more serious than those which struck Estonia in 2007. Whereas the methodology of DDoS attacks renders them non-invasive operations, State B's cyberattacks here specifically penetrate A's digital defenses in order to compromise sensitive systems directly. Likewise, while the Estonian disruptions were largely confined to the digital sphere, the attacks against State A have more material consequences in the form of displaced migrants. Despite this, the attacks do not aim to inflict direct, physical harms onto State A, instead being limited to disruption and destruction across digital systems and databases. The emphasis on digital harms renders it difficult to draw analogues between State B's cyberattacks and past uses of force given that commonly recognized uses of force have hitherto been constrained to those actions "that injure or kill persons or physically damage or destroy objects".[406] The absence of a neat analogue removes one avenue by which State B's cyber operation may come to be regarded as a use of force.

Despite this, the *Tallinn Manual* asserts that an act may nonetheless be treated as a use of force if it is widely regarded as such by other members of the international community. To this end, the *Tallinn Manual* identifies a non-exhaustive array of factors that may contribute to evaluations of cyberattacks, including attack severity, immediacy of consequences, invasiveness, and the measurability of its effects.[407] As a result, the cyber operation facing State A could be deemed a use of force should it satisfy a sufficient set of criteria. The attack can readily be considered invasive insofar as it was designed to penetrate cyber defenses in order to access the relevant systems and databases. Likewise, it may be regarded as causing immediate, potentially severe, consequences to State A's self-deliberated national interests by virtue of significantly limiting mobility within the country, as well as adding further strains in the form of a rapid influx of now-stranded migrants in need of shelter and care. The elevated financial costs associated with immediate triage and support for these migrants may also deal a significant blow to State A's economy in the short term; the *Tallinn Manual* Experts acknowledge that some states (albeit

[406] *Tallinn 2.0*, 333.
[407] *Tallinn 2.0*, 333-336.

not all) may regard economically crippling cyberattacks as constituting uses of force.[408] Considered holistically, it likely that State B's cyber operation might motivate *some* states to regard it as a use of force, potentially justifying a multilateral response in accordance with Articles 41 and 42.

While it is possible that a multilateral intervention may eventually be motivated by a cyber operation, there remains a key pragmatic concern. Multilateral responses are contingent on first having achieved consensus amongst the members of the responding international organization. The diplomatic processes necessary for forming consensus are often long-winded, requiring prolonged periods of deliberation, thereby reducing the likelihood of a timely response in the immediate aftermath of a severe cyberattack. These already-expected diplomatic delays are further exacerbated by the higher likelihood of subjective disagreements which accompanies operations carried out within the cyber domain. For some (*other*) states, disruptions or destruction limited to cyberspace might be regarded as insufficiently severe for declaring such operations uses of force. Likewise, the standards for which kinds of attacks qualify as "invasive" may vary wildly, potentially precluding cyberattacks against states with weak or nonexistent defensive cyber capabilities from ever being sufficiently invasive. These differences of opinion may serve to stall multilateral deliberations, affording an aggressor a greater window of time for continuing its cyber operation unabated. While the necessary consensus may eventually be reached, the lack of a standardized set of criteria for evaluating disanalogous uses of force ensures that most multilateral responses to cyber operations are likely to be delayed.

In the interests of motivating a swift response to State B's cyber operation and minimizing the harms it is projected to inflict, State A may seek to invoke the inherent right to self-defense outlined within Article 51. However, for State A to be morally justified in responding unilaterally with force, the scale and effects of State B's attack must first rise to that of an armed attack. Unfortunately for State A, State B's cyber operation falls short of meeting this threshold under the LOAC approach. While cyberattacks inflicting significant physical harms constitute armed attacks, the *Tallinn Manual* specifies that cyber operations "that involve brief or periodic interruption of non-essential cyber services, do not qualify as armed attacks";[409] as such, the disruptions caused by State B's cyber efforts would fail to garner the requisite

[408] *Tallinn 2.0*, 336.
[409] *Tallinn 2.0*, 341.

degree of attention. State B's cyber operation fails to directly inflict severe injury to multiple people or cause widespread physical destruction. It is possible that the cyberattack may *indirectly* result in harms trending in this direction, such as mass injuries or large-scale physical destruction stemming from riots breaking out within State A due to the sabotaged migration initiative. However, despite arguments in favour of escalating these cyberattacks to the level of armed attacks, this position does not receive unanimous support amongst the *Tallinn Manual* Experts. Within discussions pertaining to Rule 72, the Experts note internal division regarding "whether the effects in question must have been intended", with a minority of the Experts asserting that unintended consequences stemming from cyber operations are insufficient for escalating an operation to the level of an armed attack.[410] Insofar as the Experts were unable to form a consensus on this front, it remains highly likely that any forceful response deployed on these grounds by State A would be regarded by the international community as controversial.

As a result, the LOAC framework is unlikely to regard State B's cyber operation as having uncontroversially satisfied the just cause criterion, leaving State A at a veritable disadvantage. In the absence of just cause, permissible responses at State A's disposal are limited to non-forceful means, such as diplomatic and economic sanctions. State B's attacks highlight the shortcomings of tight-to-the-law approaches to the evaluation of cyber operations resulting in non-physical direct harms. Within the cyber domain, this would largely restrict State A's options to defensive cyber measures which have hitherto proven to pale in effectiveness when compared to their offensive kin. Barring the presence of widespread physical destruction or significant bodily harm and/or death, it remains highly unlikely that a cyberattack will be considered "armed" by a strictly legal approach such as that of the *Tallinn Manual*, thereby limiting the ability of states to invoke Article 51. Likewise, without analogous past cases, evaluations of the severity of cyberattacks become reliant on subjective analyses of variable criteria. This renders the process of garnering the sort of consensus necessary for motivating multilateral responses much more arduous, as states may regard the methodology and effects of the same cyberattacks in disparate lights, a problem magnified by the comparatively unsettled nature of activities within cyberspace. As such, even multilateral responses to cyber operations resulting in solely non-physical harms prove difficult to motivate under tight-to-the-law evaluative frameworks.

[410] *Tallinn 2.0*, 341.

In sharp contrast, the normative approach offered by a cyber *jus ad bellum* framework proves capable of encompassing these kinds of serious, albeit non-physically destructive, cyber operations. Assuming State A's actions genuinely reflect the self-deliberations of its citizens, State B's cyberattacks represent an impermissible violation of State A's political autonomy. State A has freely chosen to help the migrants from neighbouring states escape climate crisis: something it has every sovereign right to do. By compromising various layers of State A's digital infrastructure to weaken its ability to effectively process migrants, State B directly undermines State A's political self-deliberations and instead tries to ensure a specific politically relevant outcome aligned with its own interests. Provided the attack is designed to penetrate State A's cyber defenses, rather than solely exploiting unguarded networks and infrastructure as is the case in DDoS attacks, it satisfies the just cause criterion.[411] As a result, unlike within the LOAC framework, this type of cyber operation would be regarded as sufficient just cause for forceful unilateral responses in the interests of self-defense.

Despite unilateral force being a theoretically permissible response in accordance with this criterion, many of the same considerations within the Stuxnet case continue to apply. A forceful *kinetic* response nonetheless remains an unlikely option by virtue of the constraints imposed by the remaining *jus ad bellum* criteria. The direct damages wrought by State B's cyber operation are digital in nature. As such, it remains exceedingly difficult to compare their damages with those inflicted by conventional armed attacks without wading into controversial waters; the deliberations required to navigate these kinds of comparisons would inherently delay the immediacy with which a forceful response may be deployed, thereby precluding a state's ability to respond to cyber aggression in a timely fashion with kinetic measures. Consequently, kinetic responses to State B's cyber operations are highly likely to violate the principle of proportionality. By contrast, cyber retaliations are capable of inflicting harms readily comparable to those incurred by State B's initial act of cyber aggression. For example, State A may deploy disruptive cyberattacks of its own against infrastructure located within State B, inflicting a punitive cost on State B with more immediacy than would be feasibly achievable via diplomatic or economic measures. Responding to cyber aggression with cyber retaliation reduces the risks of a grossly disproportionate response stemming from subjective comparisons of disparate kinds

[411] Smith, "Cyberattacks as Casus Belli: A Sovereignty-Based Account", 231.

of attacks, which may in turn motivate an escalation to kinetic responses or, in extreme cases, full-blown conventional warfare.

Further limitations on potential responses are imposed by the principle of reasonable probability of success. It is difficult to formulate a feasible kinetic response available to State A that enjoys a high likelihood of halting further cyber aggression by State B, while still adhering to the remaining principles of *jus ad bellum*. A deterrent deployment of kinetic force designed to inflict an unignorable cost on State B would presumably be effective, however such a response is highly likely to violate the proportionality principle in all but the most extreme of cases within which the initial cyberattack inflicted significant physical harms. Responses of this nature would likewise draw great scrutiny regarding the intentions underpinning State A's response insofar as a disproportionately harmful retaliation suggests State A's motivations go beyond merely defending against aggression, and instead veer into the territory of revenge. The less-inherently-harmful nature of cyber responses allows them to satisfy the principle of right intention more readily by offering evidence of a commitment on State A's behalf to avoid inflicting unnecessary harms. By willfully constraining its response to non-physically destructive cyberattacks, State A offers support for its assertions that its response is motivated solely by the desire to defend itself against foreign aggression, rather than seizing the opportunity to advance further secretive foreign policy objectives.

Yet another restraint on State A's ability to respond to cyber aggression with kinetic force emerges in the form of the principle of last resort. In the event of diplomatic and economic efforts at curbing State B's aggression proving ineffective, State A remains unjustified in immediately escalating to kinetic measures. Cyber operations serve to bridge the gap between stronger economic measures and weaker kinetic attacks, offering states a broader spectrum of progressively escalatory responses. As such, the principle of last resort demands that State A first explore feasible cyber alternatives capable of halting State B's cyber operation. For example, State A may discover that a logic bomb targeting State B's cyber-specific intelligence apparatus is likely to significantly hinder State B's cyber operation, if not bring it to a halt entirely. While the primary concern here is the immediate moral right of State A to respond to cyber aggression with force, it is worth noting that allies of State A may justifiably intervene on State A's behalf once it possesses just cause, in accordance with international collective security agreements. Accordingly, even if a viable cyber response is unavailable to State A due to it possessing only

rudimentary cyber capabilities, one of its more cyber-capable allies may nonetheless have transitive permission to deploy cyber measures against State B on its behalf to defend it from foreign aggression. The proliferation of less-forceful cyber alternatives, whether conducted by State A itself or on its behalf by an allied nation, renders it more likely that an effective response to B's cyber aggression can be found without having to explore kinetic options.

Accordingly, while tight-to-the-law evaluative approaches restrict the ability of states to invoke the right to self-defense in the event of non-physically destructive cyberattacks, the cyber *jus ad bellum* framework acknowledges that some such attacks may morally justify forceful responses. While the *jus ad bellum* framework requirements for just cause render it more likely that a cyberattack may fulfill that criterion, the subsequent principles nonetheless limit the kinds of forceful responses that states are morally justified in deploying. Despite these limitations largely precluding the justifiable deployment of kinetic responses to cyber operations, the *jus ad bellum* approach nonetheless offers states the ability to justifiably defend themselves with *cyber force in kind*. Although the LOAC approach limits permissible responses to short-of-force measures, this alternative framework preserves the inherent right to forceful self-defense within cyberspace even when the harm being incurred is digital and non-physically destructive in nature. This preservation of the right to self-defense in the face of cyberattacks mitigates the likelihood of an escalation of hostilities by limiting permissible forceful responses to solely cyber measures in all but the most extreme of scenarios, within which a kinetic response would likely be justified even in accordance with the more restrictive LOAC approach.

7.5 A More Flexible Approach to Cyberwar Governance

In each of the three above cases, the cyber *jus ad bellum* framework offers a flexible evaluative approach which is better equipped for accommodating a fuller spectrum of cyberattacks than current tight-to-the-law frameworks. While the tight-to-the-law approach runs into controversy regarding the severity of the Stuxnet worm, the cyber *jus ad bellum* framework asserts that it satisfies the just cause criterion, motivating potentially forceful responses, but nonetheless limiting morally permissible responses to cyber uses of force to adhere to the remaining principles of *jus ad bellum*. In the case of Russia's cyber campaigns in Crimea and the

Donbas, the cyber *jus ad bellum* approach identifies the acute threat cyber interference and information control poses to the security of the state. Unlike in the LOAC approach within which forceful responses only become permissible once the cyber effort evolves into hybrid warfare, the *jus ad bellum* approach permits more proactive cyber responses to tackle offensive cyber operations before they achieve their full objectives. Finally, in the hypothetical case resulting in no direct physical harms, the cyber *jus ad bellum* approach offers provisions for evaluating cyberattacks resulting in non-tangible harms as nonetheless potentially constituting just cause. While states may be justified in deploying force in response to these cyberattacks, they remain highly unlikely to be morally justified in utilizing anything but non-physically harmful cyberattacks in response, due to the constraints imposed by the remaining *jus ad bellum* criteria.

Chapter 8

Conclusion

8.1 Navigating the Digitization of Warfare

The paradigmatic shift away from conventional interstate conflict has left us in a lurch. Advancements in network technologies have led to rapid digitization across all levels of modern society. Concurrently, state military and intelligence organizations have ramped up the development of sophisticated cyberweapons capable of exploiting emerging vulnerabilities which accompany this transition. As the frequency, reach, and severity of cyberattacks continues to rise, the absence of a governing framework grows more and more apparent. In the interim, states deploy cyberweapons in a trial-and-error fashion, probing the capabilities and tolerances of other states, unfettered by the LOAC which constrain their activities within the traditional domains of war. I have argued that the resulting status quo is unacceptably precarious, as the lack of defined thresholds or norms of governance renders it possible for cyber-capable states to grossly overstep an undefined boundary and trigger an otherwise avoidable escalation of hostilities.

Although the *Tallinn Manual* offers a strong initial foray into cyberwar governance, its strict adherence to the existing LOAC ultimately limits its evaluative capabilities within cyberspace. Both the *Tallinn Manual*'s conceptions of use of force and armed attack are heavily reliant on the presence of *kinetic damages*; cyberattacks must be destructive within *real space* in order to motivate forceful responses. However, the vast majority of cyber operations fail to clear this lofty standard, instead prioritizing disruptions of physical and digital infrastructure, as well as the destruction of digital assets. As a result, aggressor states largely operate with impunity under this tight-to-the-law approach, secure in the knowledge that cyber operations failing to inflict kinetic-equivalent harms will likewise fail to motivate forceful reprisals. The analogy approach of this tight-to-the-law evaluative framework serves to further compound the disadvantages facing cyber defenders.

This project seeks to address this imbalance by preserving the inherent right of states to self-defense, even in the absence of glaring kinetic harms. In pursuit of this objective, this project

has turned towards the *jus ad bellum* of JWT. By broadening the *just cause* criterion to encompass the novel risks cyber warfare poses to the state, I have identified the circumstances within which a disanalogous cyberattack may nonetheless pose an existential risk to the state severe enough to warrant a forceful response. Despite rendering it more likely that the just cause criterion can be met by cyber means, the remaining five criteria serve to constrain a state's ability to respond to cyber aggression with kinetic force. Under the cyber *jus ad bellum*, most incidences of incurred cyber force will motivate only cyber responses in kind, barring extreme cases within which a kinetic response is likely to be morally justified even in accordance with the more restrictive LOAC evaluative framework.

Although the cyber *jus ad bellum* is not itself legally binding, I believe that it offers a strong conceptual and ethical foundation for the evaluation of a fuller spectrum of cyber operations and the development of norms of best practice for state conduct within cyberspace. By forgoing a tight-to-the-law approach, the cyber *jus ad bellum* avoids becoming mired in controversy stemming from competing interpretations of codified law. Instead, by appealing to the interests of all relevant stakeholders, the cyber *jus ad bellum* follows in the footsteps of the *Montreux Document* as a soft law approach to navigating the current uncertainty of cyber warfare. Not only does this approach equip states with a comprehensive framework for evaluating a full range of present-day cyber operations, it also offers guidance for novel future cyberattacks for which there may be no applicable laws nor legal precedents. While the codification of law will always be outpaced by the development of new cyberweapons, a cyber *jus ad bellum* framework can go a long way towards redressing the current perilous "equilibrium" within which we find ourselves.

Bibliography

Books

Clarke, Richard A., and Knake, Robert, *Cyber War: The Next Threat to National Security and What to Do About It* (New York NY: Ecco, 2012).

Gibbs, David N. *First Do No Harm: Humanitarian Intervention and the Destruction of Yugoslavia.* (Nashville: Vanderbilt University Press, 2009).

Greenberg, Andy, *Sandworm: A New Era of Cyberwar and the Hunt for the Kremlin's Most Dangerous Hackers*. (New York, NY: Doubleday, 2019).

Jamieson, Kathleen Hall, *Cyberwar: How Russian Hackers and Trolls Helped Elect a President* (New York, NY: Oxford University Press, 2018).

Jasper, Scott, *Russian Cyber Operations: Coding the Boundaries of Conflict*, (Washington, DC: Georgetown University Press, 2020).

Lucas, George R. *Ethics and Cyber Warfare: The Quest for Responsible Security in the Age of Digital Warfare*. (New York, NY: Oxford University Press, 2017).

Nettelfield, Lara J., and Wagner, Sarah E., *Srebrenica in the Aftermath of Genocide* (Cambridge; New York: Cambridge University Press, 2014).

Orend, Brian, *Michael Walzer on War and Justice*, (Montreal; Ithaca, [N.Y.]: McGill-Queen's University Press, 2000).

Orend, Brian, *The Morality of War*, (Peterborough, Ont.: Broadview Press, 2006).

Orend, Brian, *War and Political Theory.* (Cambridge, UK; Medford, MA: Polity, 2019).

Roudometof, Victor, and Robertson, Roland, *Nationalism, Globalization, and Orthodoxy: The Social Origins of Ethnic Conflict in the Balkans*, Contributions to the Study of World History, no. 89 (Westport, CT: Greenwood Press, 2001).

Schmitt, Michael N., and NATO Cooperative Cyber Defence Centre of Excellence, eds. *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations*. Second edition. (Cambridge, United Kingdom; New York, NY, USA: Cambridge University Press, 2017).

Walzer, Michael, *Just and Unjust Wars: A Moral Argument with Historical Illustrations,* 4th ed. (New York: Basic Books, 2006).

Walzer, Michael, *Spheres of Justice: A Defense of Pluralism and Equality*. (New York: Basic Books, 2010).

Book Chapters

Acton, James M., "Cyber Weapons and Precision-Guided Munitions," in *Understanding Cyber Conflict: Fourteen Analogies.* Eds. Perkovich, George, and Ariel E. Levite. (Washington: Georgetown University Press, 2017). 45-60. muse.jhu.edu/book/62546.

Danchev, Alex, and Keohane, Dan, "Introduction: The Rules of Propriety," in *International Perspectives on the Gulf Conflict, 1990-91*, eds. Alex Danchev & Dan Keohane. (New York: St. Martin's Press in association with St. Antony's College, Oxford, 1994).

Dipert, Randall R., "Distinctive Ethical Issues of Cyberwarfare," In *Binary Bullets: The Ethics of Cyberwarfare*, eds. Fritz Allhoff, Adam Henschke, & Bradley Jay Strawser (New York, NY: Oxford University Press, 2016).

Foot, M. R. D., "Conditions Making For Success and Failure of Denial and Deception: Democratic Regimes", In *Denial and Deception: The Twenty-First Century Challenge*, eds. Roy Godson & James J. Wirtz (New Brunswick, N.J: Transaction Publishers, 2002).

Jenkins, Ryan, "Cyberwarfare as Ideal War," in *Binary Bullets: The Ethics of Cyberwarfare*, eds. Fritz Allhoff, Adam Henschke, & Bradley Jay Strawser (New York, NY: Oxford University Press, 2016).

Lee, Steven P., "The Ethics of Cyberattack," In *The Ethics of Information Warfare*, eds. Luciano Floridi & Mariarosaria Taddeo (Heidelberg: Springer International Publishing, 2014).

Lemarchand, René, "The 1994 Rwanda Genocide," in *Century of Genocide: Critical Essays and Eyewitness Accounts*, eds. 3rd ed. Samuel Totten and William S. Parsons, (New York: Routledge, 2009).

Lin, Patrick, Fritz Allhoff, and Keith Abney, "Is Warfare the Right Frame for the Cyber Debate?", In *The Ethics of Information Warfare*, eds. Luciano Floridi & Mariarosaria Taddeo (Heidelberg: Springer International Publishing, 2014).

Lucas, George R., "Emerging Norms for Cyber Warfare," In *Binary Bullets: The Ethics of Cyberwarfare*, eds. Fritz Allhoff, Adam Henschke, & Bradley Jay Strawser (New York, NY: Oxford University Press, 2016).

Orend, Brian, "Fog in the Fifth Dimension: The Ethics of Cyber-War," in *The Ethics of Information Warfare*, eds. Luciano Floridi and Mariarosaria Taddeo, vol. 14, Law, Governance and Technology Series (Cham: Springer International Publishing, 2014), 3–23, https://doi.org/10.1007/978-3-319-04135-3_1.

Pakharenko, Glib, "Cyber Operations at Maidan: A First-Hand Account," in in *Cyber War in Perspective: Russian Aggression against Ukraine*, ed. Kenneth Geers (NATO Cooperative Cyber Defence Centre of Excellence, 2015).

https://www.ccdcoe.org/library/publications/cyber-war-in-perspective-russian-aggression-against-ukraine/.

Schmitt, Michael N., and Vihul, Liis, "The Emergence of International Legal Norms for Cyberconflict," in *Binary Bullets: The Ethics of Cyberwarfare*, eds. Fritz Allhoff, Adam Henschke, & Bradley Jay Strawser (New York, NY: Oxford University Press, 2016).

Stinissen, Jan, "A Legal Framework for Cyber Operations in Ukraine," in *Cyber War in Perspective: Russian Aggression against Ukraine*, ed. Kenneth Geers (NATO Cooperative Cyber Defence Centre of Excellence, 2015).

https://www.ccdcoe.org/library/publications/cyber-war-in-perspective-russian-aggression-against-ukraine/.

Whetham, Davidm "Cyber *Chevauchées*," In *Binary Bullets: The Ethics of Cyberwarfare*, eds. Fritz Allhoff, Adam Henschke, & Bradley Jay Strawser (New York, NY: Oxford University Press, 2016).

Journal Articles

Allen, T. S., and Moore, A. J., "Victory without Casualties: Russia's Information Operations," *Parameters* 48, no. 1 (2018), https://press.armywarcollege.edu/parameters/vol48/iss1/8.

Barrett, Edward T., "Warfare in a New Domain: The Ethics of Military Cyber-Operations," *Journal of Military Ethics* 12, no. 1 (April 2013): 4–17. https://doi.org/10.1080/15027570.2013.782633.

Bass, Gary J., "Jus Post Bellum", *Philosophy & Public Affairs* 32, no. 4 (2004): 384–412. http://www.jstor.org/stable/3557994.

Biersack, John, and O'Lear, Shannon, "The Geopolitics of Russia's Annexation of Crimea: Narratives, Identity, Silences, and Energy," *Eurasian Geography and Economics* Vol. 55. Issue 3. (December 2014), accessed July 11, 2021: https://doi.org/10.1080/15387216.2014.985241.

Boylan, Eric, "Applying the Law of Proportionality to Cyber Conflict: Suggestions for Practitioners," *Vanderbilt Journal of Transnational Law*, vol. 50, no. 1 (January 2017): 217-246.

Brown, Gary D., "Proportionality and Just War," *Journal of Military Ethics* 2, no. 3 (November 2003): 171–85. https://doi.org/10.1080/15027570310000667.

Condra, Luke N., and Shapiro, Jacob N., "Who Takes the Blame? The Strategic Effects of Collateral Damage," *American Journal of Political Science* 56, no. 1 (January 2012): 167–87. https://doi.org/10.1111/j.1540-5907.2011.00542.x.

Cormac, Rory, and Aldrich, Richard J., "Grey Is the New Black: Covert Action and Implausible Deniability," *International Affairs* 94, no. 3 (May 1, 2018): 477–494. https://doi.org/10.1093/ia/iiy067.

Farwell, James P., and Rohozinski, Rafal, "Stuxnet and the Future of Cyber War," *Survival*, Vol. 53. Issue 1. (January 2011), accessed May 27, 2021: https://doi.org/10.1080/00396338.2011.555586.

Fridman, Ofer, "Information War as the Russian Conceptualisation of Strategic Communications," *The RUSI Journal* Vol. 165. Issue 1. (March 2020), accessed June 5, 2021: https://doi.org/10.1080/03071847.2020.1740494.

Galtung, Fredrik, and Tisné, Martin, "A New Approach to Postwar Reconstruction," *Journal of Democracy* 20, no. 4 (2009): 93–107. https://doi.org/10.1353/jod.0.0132.

Ghafur, S., Kristensen, S., Honeyford, K. *et al.* A retrospective impact analysis of the WannaCry cyberattack on the NHS. *npj Digit. Med.* 2, 98 (2019). https://doi.org/10.1038/s41746-019-0161-6.

Giuliano, Elise, "Who Supported Separatism in Donbas? Ethnicity and Popular Opinion at the Start of the Ukraine Crisis," *Post-Soviet Affairs* 34, no. 2–3, May 4, 2018: 158–78. https://doi.org/10.1080/1060586X.2018.1447769.

Greenwood, Christopher, "New World Order or Old? The Invasion of Kuwait and the Rule of Law," *The Modern Law Review* Vol. 55, Issue 2 (March 1992), accessed July 2, 2021: https://doi.org/10.1111/j.1468-2230.1992.tb01870.x.

Helm, Toby, Luke Harding, Daniel Boffey, and Julian Borger, "Defiant Putin Warns the West: Your Sanctions are Akin to an Act of War," *The Guardian*, March 5, 2022. Accessed March 21, 2022.

https://www.theguardian.com/world/2022/mar/05/defiant-putin-warns-the-west-your-sanctions-are-akin-to-an-act-of-war.

Herring, George C., "America and Vietnam: The Unending War," *Foreign Affairs* 70, no. 5 (1991): 104-119. https://doi.org/10.2307/20045006.

Iasiello, Emilio J, "Russia's Improved Information Operations: From Georgia to Crimea." *The US Army War College Quarterly: Parameters* 47, no. 2 (2017).

Jensen, Benjamin, Brandon Valeriano, and Ryan Maness, "Fancy Bears and Digital Trolls: Cyber Strategy with a Russian Twist," *Journal of Strategic Studies* Vol. 42. Issue 2. (January 2019), accessed May 29, 2021: https://doi.org/10.1080/01402390.2018.1559152.

Jervis, Robert, “Reports, Politics, and Intelligence Failures: The Case of Iraq,” *Journal of Strategic Studies* 29, no. 1 (February 2006): 3–52. https://doi.org/10.1080/01402390600566282.

Kilovaty, Ido, “Cyber Warfare and the Jus Ad Bellum Challenges: Evaluation in the Light of the Tallinn Manual on the International Law Applicable to Cyber Warfare,” *National Security Law Brief* 5, no. 5 (2014): 90-124.

Koeman, Annalisa, “A Realistic and Effective Constraint on the Resort to Force? Pre-Commitment to *Jus in Bello* and *Jus Post Bellum* as Part of the Criterion of Right Intention”, *Journal of Military Ethics* 6, no. 3 (September 2007): 198–220. https://doi.org/10.1080/15027570701585373.

Langner, Ralph, "Stuxnet: Dissecting a Cyberwarfare Weapon," in *IEEE Security & Privacy*, vol. 9, no. 3, pp. 49-51. May-June 2011, doi: 10.1109/MSP.2011.67.

Leonhardt, David, and Philbrick, Ian Prasad, “*Economic War*,” *The New York Times*, March 11, 2022. Accessed March 21, 2022. https://www.nytimes.com/2022/03/11/briefing/economic-war-sanctions-russia.html.

Lindsay, Jon R., “Stuxnet and the Limits of Cyber Warfare.” *Security Studies* 22, no. 3 (July 1, 2013): 365–404. doi:10.1080/09636412.2013.816122.

Lippman, Matthew, "War Crimes: The My Lai Massacre and the Vietnam War," San Diego Justice Journal 1, no. 2 (Summer 1993): 295-364. https://heinonline.org/HOL/P?h=hein.journals/tjeflr15&i=301.

McKay, Spencer, and Tenove, Chris, “Disinformation as a Threat to Deliberative Democracy,” *Political Research Quarterly* 74, no. 3 (September 2021): 703–17. https://doi.org/10.1177/1065912920938143.

O’Brien, Kevin A., “Interfering with Civil Society: CIA and KGB Covert Political Action during the Cold War,” *International Journal of Intelligence and Counterintelligence* 8, no. 4 (December 1995): 431–56. https://doi.org/10.1080/08850609508435297.

Orend, Brian, “Justice after War,” *Ethics & International Affairs* 16, no. 1 (March 2002): 43–56. https://doi.org/10.1111/j.1747-7093.2002.tb00374.x.

Orend, Brian, "Kant's Just War Theory." *Journal of the History of Philosophy* 37, no. 2 (Apr 01, 1999): 323.

http://search.proquest.com.proxy.lib.uwaterloo.ca/scholarly-journals/kants-just-war-theory/docview/1297333042/se-2?accountid=14906.

Orend, Brian, “Michael Walzer on Resorting to Force,” *Canadian Journal of Political Science*, Vol. 33. Issue 3. (September 2000).

Parmar, Inderjeet, “Catalysing Events, Think Tanks and American Foreign Policy Shifts: A Comparative Analysis of the Impacts of Pearl Harbor 1941 and 11 September 2001,” *Government and Opposition* 40, no. 1 (2005): 1–25. https://doi.org/10.1111/j.1477-7053.2005.00141.x.

Peisakhin, Leonid, and Rozenas, Arturas, “Electoral Effects of Biased Media: Russian Television in Ukraine,” *American Journal of Political Science*, 62(3), January 19, 2018: 535-550. http://dx.doi.org/10.2139/ssrn.2937366.

Rid, Thomas, “Cyber War Will Not Take Place,” *Journal of Strategic Studies* 35, no. 1 (February 2012): 5–32. https://doi.org/10.1080/01402390.2011.608939.

Schulzke, Marcus, “The Politics of Attributing Blame for Cyberattacks and the Costs of Uncertainty,” *Perspectives on Politics* 16, no. 4 (December 2018): 954–68. https://doi.org/10.1017/S153759271800110X.

Sherwin, Martin J. "Hiroshima as Politics and History." *The Journal of American History* 82, no. 3 (1995): 1085–93. https://doi.org/10.2307/2945113.

Smith, Patrick Taylor, "Cyberattacks as Casus Belli: A Sovereignty-Based Account," *Journal of Applied Philosophy* 35, no. 2 (May 2018): 222–41. https://doi.org/10.1111/japp.12169.

Souleimanov, Emil Aslan, Eduard Abrahamyan, and Huseyn Aliyev, "Unrecognized States as a Means of Coercive Diplomacy? Assessing the Role of Abkhazia and South Ossetia in Russia's Foreign Policy in the South Caucasus," *Southeast European and Black Sea Studies* 18, no. 1 (January 2, 2018): 73–86. https://doi.org/10.1080/14683857.2017.1390830.

Suchman, Edward A., Rose K. Goldsen, and Robin M. Williams, "Attitudes Toward the Korean War." *The Public Opinion Quarterly* 17, no. 2 (1953): 171–84. http://www.jstor.org/stable/2746273.

Tong, Chau, et al., "'Fake News Is Anything They Say!' — Conceptualization and Weaponization of Fake News among the American Public," *Mass Communication and Society* 23, no. 5 (September 2, 2020): 755–78. https://doi.org/10.1080/15205436.2020.1789661.

Wilner, Alex S., "US Cyber Deterrence: Practice Guiding Theory," *Journal of Strategic Studies* Vol. 43. Issue 2. (February 2019), accessed June 5, 2021. https://doi.org/10.1080/01402390.2018.1563779.

News Articles

Borger, Julian, and Chulov, Martin, "US kills Iran General Qassem Suleimani in strike ordered by Trump," *The Guardian*, January 3, 2020. Accessed March 22, 2022.

https://www.theguardian.com/world/2020/jan/03/baghdad-airport-iraq-attack-deaths-iran-us-tensions.

Borger, Julian, and Watt, Nicholas, "G7 Countries Snub Putin and Refuse to Attend Planned G8 Summit in Russia," *The Guardian*, March 24, 2014. Accessed March 21, 2022. https://www.theguardian.com/world/2014/mar/24/g7-countries-snub-putin-refuse-attend-g8-summit-russia.

Borger, Julian, and Wintour, Patrick, "Missiles launched by Iran against US airbases in Iraq," *The Guardian*, January 8, 2020. Accessed March 31, 2022. https://www.theguardian.com/world/2020/jan/07/trump-iran-suleimani-threats-retaliation.

Bumiller, Elisabeth, and Shanker, Thom, "Panetta Warns of Dire Threat of Cyberattack on U.S.," *The New York Times*, October 11, 2012. Accessed January 15, 2022. https://www.nytimes.com/2012/10/12/world/panetta-warns-of-dire-threat-of-cyberattack.html.

Cotovio, Vasco, and Alberti, Mia, "Anonymous claims responsibility for "ongoing" hacking of Russian government sites," *CNN*, February 26, 2022. Accessed March 7, 2022.

https://www.cnn.com/europe/live-news/ukraine-russia-news-02-26-22/h_b7cb924af9172ecfce000603d4c8b5e9.

Deutsche Welle, "Rural German District Declares Disaster After Cyberattack", July 10, 2021, Accessed August 8, 2021. https://www.dw.com/en/rural-german-district-declares-disaster-after-cyberattack/a-58227484.

Dilanian, Ken, "Code in Huge Ransomware Attack Written to Avoid Computers That Use Russian, Says New Report," *NBC News*, July 7, 2021, accessed July 8, 2021.

https://www.nbcnews.com/politics/national-security/code-huge-ransomware-attack-written-avoid-computers-use-russian-says-n1273222.

Fung, Brian, and Cohen, Zachary, "Russian Military Targeted Passwords in Wide-Ranging Hacking Campaign, US and UK Officials Say," *CNN*, July 1, 2021, accessed July 8, 2021.

https://www.cnn.com/2021/07/01/politics/russian-military-hacking-campaign-us-uk-advisory/index.html.

Gonzalez, Gloria, Ben Lefebvre, and Eric Geller, "Jugular of the U.S. Fuel Pipeline System Shuts Down After Cyberattack," *Politico*, May 8, 2021, accessed May 23, 2021. https://www.politico.com/news/2021/05/08/colonial-pipeline-cyber-attack-485984.

Goodin, Dan. "Florida Water Plant Compromise Came Hours After Worker Visited Malicious Site," *Ars Technica*, May 18, 2021, accessed May 23, 2021.

https://arstechnica.com/gadgets/2021/05/florida-water-plant-compromise-came-hours-after-worker-visited-malicious-site/.

Kramer, Andrew E., "Ukraine Cyberattack Was Meant to Paralyze, Not Profit", *The New York Times*. June 28, 2017, Accessed September 20, 2021.

https://www.nytimes.com/2017/06/28/world/europe/ukraine-ransomware-cyberbomb-accountants-russia.html.

Kuo, Lily, and Kommenda, Niko, "What is China's Belt and Road Initiative?", *The Guardian*, July 30, 2018, accessed July 7, 2021.

https://www.theguardian.com/cities/ng-interactive/2018/jul/30/what-china-belt-road-initiative-silk-road-explainer.

Dan Lamothe, "These U.S. troops survived one of the greatest crises of the Trump era. A year later, they're still coping." *The Washington Post*, January 10, 2021. Accessed March 31, 2022.

https://www.washingtonpost.com/national-security/us-military-iran-missile-attack/2021/01/10/651c3930-4fb0-11eb-b2e8-3339e73d9da2_story.html.

Medbury, Jennifer, and Haskell-Dowland, Paul, "The hacker group Anonymous has waged a cyber war against Russia. How effective could they actually be?" *The Conversation*, February 28, 2022. Accessed March 7, 2022.

https://theconversation.com/the-hacker-group-anonymous-has-waged-a-cyber-war-against-russia-how-effective-could-they-actually-be-178034.

Milmo, Dan, "Anonymous: the hacker collective that has declared cyberwar on Russia," *The Guardian*, February 27, 2022. Accessed March 7, 2022.

https://www.theguardian.com/world/2022/feb/27/anonymous-the-hacker-collective-that-has-declared-cyberwar-on-russia.

Njini, Felix, "Crippled Africa Port Goes Manual as PE-Backed Software is Shut", *Bloomberg*. July 22, 2021. Accessed September 24, 2021.

https://www.bloomberg.com/news/articles/2021-07-22/south-africa-s-transnet-reports-disruption-to-it-services.

Paul, Kari, "Who's behind the Kaseya ransomware attack – and why is it so dangerous?" *The Guardian*, July 7, 2021. Accessed March 7, 2022.

https://www.theguardian.com/technology/2021/jul/06/kaseya-ransomware-attack-explained-russia-hackers.

Sanger, David E., and Perlroth, Nicole, "*U.S. Escalates Online Attacks on Russia's Power Grid,*" *The New York Times*, June 15, 2019, accessed July 8, 2021. https://www.nytimes.com/2019/06/15/us/politics/trump-cyber-russia-grid.html.

Sanger, David E., Nicole Perlroth, and Eric Schmitt, "Scope of Russian Hack Becomes Clear: Multiple U.S. Agencies Were Hit," *New York Times*, December 14, 2020, accessed May 23, 2021.

https://www.nytimes.com/2020/12/14/us/politics/russia-hack-nsa-homeland-security-pentagon.html.

Schectman, Joel, Christopher Bing, and James Pearson, "Ukrainian cyber resistance group targets Russian power grid, railways," *Reuters*, March 1, 2022. Accessed March 7, 2022.

https://www.reuters.com/technology/ukrainian-cyber-resistance-group-targets-russian-power-grid-railways-2022-03-01/.

Sengupta, Somini, and Kramer, Andrew K. "Dutch Inquiry Links Russia to 298 Deaths in Explosion of Jetliner Over Ukraine," *The New York Times*. September 28, 2016. Accessed February 16, 2022.

https://www.nytimes.com/2016/09/29/world/asia/malaysia-air-flight-mh17-russia-ukraine-missile.html.

Shane, Scott, "*Drone Strikes Reveal Uncomfortable Truth: U.S. Is Often Unsure About Who Will Die*," *The New York Times*, April 23, 2015. Accessed March 15, 2022.

https://www.nytimes.com/2015/04/24/world/asia/drone-strikes-reveal-uncomfortable-truth-us-is-often-unsure-about-who-will-die.html.

Smith, R. Jeffrey, "The Failed Reconstruction of Iraq," *The Atlantic*, March 15, 2013. Accessed February 14, 2022.

https://www.theatlantic.com/international/archive/2013/03/the-failed-reconstruction-of-iraq/274041/.

Suliman, Adela, "Sweden sets up Psychological Defense Agency to fight fake news, foreign interference," *The Washington Post*, January 6, 2022. Accessed January 16, 2022.

https://www.washingtonpost.com/world/2022/01/06/sweden-fake-news-psychological-defence-agency/.

Tapper, Jake, Evan Perez, and Ryan Young, "Cox Media Group hit by cyberattack last week, sources familiar tell CNN," *CNN*, June 9, 2021. Accessed January 16, 2022. https://www.cnn.com/2021/06/09/politics/cox-media-group-cyberattack/index.html.

Troianovski, Anton, "Moscow orders troops to Ukraine's separatist regions after Putin recognizes their independence," *The New York Times.* February 21, 2022. Accessed February 23, 2022.

https://www.nytimes.com/live/2022/02/21/world/ukraine-russia-putin-biden/moscow-orders-troops-to-ukraines-breakaway-regions-for-peacekeeping-functions.

Troianovski, Anton, Roger Cohen, and Katie Rogers, "Putin Warns the West and Ukraine but Keeps His Intentions a Mystery," *The New York Times*, February 7, 2022. Accessed February 9, 2022. https://www.nytimes.com/2022/02/07/world/europe/putin-macron-russia-france-ukraine.html.

Troianovski, Anton, and Schwirtz, Michael, "As Russia Stalls in Ukraine, Dissent Brews Over Putin's Leadership," *The New York Times*, March 22, 2022. Accessed March 23, 2022. https://www.nytimes.com/2022/03/22/world/europe/putin-russia-military-planning.html.

Ward, Mark "William Gibson says the future is right here, right now," *BBC News*, October 12, 2010. Accessed June 28, 2022. https://www.bbc.com/news/technology-11502715.

Reports

Baezner, Marie, "Cyber and Information Warfare in the Ukrainian Conflict," *Center for Security Studies (CSS)*, *ETH Zürich,* October 2018. https://doi.org/10.3929/ETHZ-B-000321570.

Bradshaw, Samantha, Howard, Philip N., "The Global Disinformation Order: 2019 Global Inventory of Organised Social Media Manipulation" *Oxford Internet Institute*, 2019, Copyright, Fair Use, Scholarly Communication, etc. https://digitalcommons.unl.edu/scholcom/207.

Connell, Michael, and Vogler, Sarah, "Russia's Approach to Cyber Warfare," *CAN*, March 2017, accessed June 2, 2021. https://www.cna.org/cna_files/pdf/DOP-2016-U-014231-1Rev.pdf.

Hansen, Flemming Splidsboel and Dansk Institut for Internationale Studier, "Russian Hybrid Warfare: A Study of Disinformation," 2017. http://pure.diis.dk/ws/files/950041/DIIS_RP_2017_6_web.pdf.

Hemsley, Kevin E. and Fisher, Ronald E., "History of Industrial Control System Cyber Incidents," *U.S. Department of Energy*, December 31, 2018. https://doi.org/10.2172/1505628.

Kulacki, Gregory, "An Authoritative Source on China's Military Space Strategy," *Union of Concerned Scientists*, March 2014, accessed June 3, 2020. https://www.ucsusa.org/sites/default/files/2019-10/China-s-Military-Space-Strategy.pdf.

Richardson, Sophie, "China's Influence on the Global Human Rights System," *Human Rights Watch*, September 14, 2020, accessed July 10, 2021. https://www.hrw.org/news/2020/09/14/chinas-influence-global-human-rights-system.

Schneier, Bruce, and Wheeler, Tarah, "Hacked drones and busted logistics are the cyber future of warfare," *The Brookings Institution*. June 4, 2021. Accessed February 14, 2022.

https://www.brookings.edu/techstream/hacked-drones-and-busted-logistics-are-the-cyber-future-of-warfare/.

Sen, Ashish Kumar, "Iran's Growing Cyber Capabilities in a Post-Stuxnet Era," *Atlantic Council*. April 10, 2015. Accessed March 31, 2022.

https://www.atlanticcouncil.org/blogs/new-atlanticist/iran-s-growing-cyber-capabilities-in-a-post-stuxnet-era/.

Strayer, W. T., R. Walsh, C. Livadas, and D. Lapsley, "Detecting Botnets with Tight Command and Control," *Proceedings. 2006 31st IEEE Conference on Local Computer Networks*, 2006, pp. 195-202, doi:10.1109/LCN.2006.322100.

Other Documents

Britannica, T. Editors of Encyclopaedia. "Persian Gulf War." *Encyclopedia Britannica*, accessed May 6, 2021, https://www.britannica.com/event/Persian-Gulf-War.

Center for Strategic & International Studies, "Significant Cyber Incidents," accessed May 15, 2021: 55, https://www.csis.org/programs/strategic-technologies-program/significant-cyber-incidents.

FireEye, "What Is a Zero-Day Exploit?", accessed July 11, 2021. https://www.fireeye.com/current-threats/what-is-a-zero-day-exploit.html.

International Committee of the Red Cross, *The Montreux Document*, (Geneva: International Committee of the Red Cross, 2009).

International Court of Justice, *Case Concerning Military and Paramilitary Activities In and Against Nicaragua (Nicaragua v. United States of America); Merits*, 27 June 1986.

Lazar, Seth, "War", *The Stanford Encyclopedia of Philosophy* (Spring 2020 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/spr2020/entries/war/. Accessed Jan 25, 2022.

Raymond, D., G. Conti, T. Cross and R. Fanelli, "A control measure framework to limit collateral damage and propagation of cyber weapons," *2013 5th International Conference on Cyber Conflict (CYCON 2013)*, 2013, pp. 1-16.

Rowe, Neil, *Towards Reversible Cyberattacks*, Reading: Academic Conferences International Limited, 2010. http://search.proquest.com.proxy.lib.uwaterloo.ca/conference-papers-proceedings/towards-reversible-cyberattacks/docview/869507120/se-2?accountid=14906.

United Nations, *Charter of the United Nations*, October 24, 1945, 1 UNTS XVI.

United Nations General Assembly, *Definition of Aggression*, 14 December 1974, A/RES/3314.

United Nations Security Council, *Resolution 678, Iraq/Kuwait*, (29 November 1990).

The White House, Briefing Room, "The United States, Joined by Allies and Partners, Attributes Malicious Cyber Activity and Irresponsible State Behavior to the People's Republic of China," July 19, 2021. Accessed March 7, 2022. https://www.whitehouse.gov/briefing-room/statements-releases/2021/07/19/the-united-states-joined-by-allies-and-partners-attributes-malicious-cyber-activity-and-irresponsible-state-behavior-to-the-peoples-republic-of-china/.