Cyber Conflict and Just War Theory

by Nicholas Collie

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy.

October 2022

Table of Contents

Prelude

The only true basis of enduring peace is the willing cooperation of free peoples in a world in which, relieved of the menace of aggression, all may enjoy economic and social security; and that it is our intention to work together, and with other free peoples, both in war and peace, to this end.

Declaration of St. James' Palace, 1941

Much of this project has been completed in the state of 'lockdown' necessitated by a global pandemic. During that process many ethical problems, that once may have seemed abstract, were made concrete. One of those is the unpleasant calculation that balances the loss of a certain number of lives against a desired political or social outcome. The calculation is tragic in nature in that, while there may be a best solution given the circumstances, there is no solution that does not contain serious moral costs.

A similar calculation is at the centre of moderate theories of war. A moderate theory of war is one that inherently tackles this type of balance. It does not claim that the benefits of the political outcome always trump the costs. Nor does it claim that nothing can trump the moral cost of the loss of lives. Rather, in the context of war, it claims that, in certain circumstances and certain ways and subject to certain limitations, the benefits of the outcome can justify the moral cost of the path to that outcome. Defining those circumstances is the aim of moderate theories and is a complex task. It would be easier if this were a clear-cut and sterile domain, but it is already clear that both moderate theories of war and the assessment of their applicability are likely to be messy. Large-scale conflict is a complex phenomenon and has an inherently tragic aspect. Any search for what might be regarded as 'surgical precision' results in such great abstractions or simplification that the analysis becomes distanced from reality to a significant degree.

One of the famous responses (Walzer 2006b:43) in the recent philosophical discussion of the theory of war has been that a particular interpretation provides "a careful and precise account of what individual responsibility in war would be like if war were a peacetime activity." Another quotation that is used later comes from Michael Neu (2013:462) when he criticises some philosophical analyses of conflict as being concerned with "neat fictitious worlds, rather than the complex real world". The quotations are used to emphasise that extreme abstraction away from the messy nature of war and large-scale conflict is possible but not desirable. We should not necessarily be looking for simple answers because those simple answers do not adequately reflect the tragic and complex circumstances.

This project is further complicated by the fact that it is an analysis of how existing moderate theories of conflict (Just War Theory is used as a salient example) accommodate cyber conflict. So, it traverses difficult terrain for another reason. Cyber conflict is a subject that is so recent that we have limited perspective on its effects. This problem is aggravated by the fact that much previous cyber conflict has been covert in nature and an extensive and authoritative history does not yet exist.

We are in a transitional time. The digital revolution is effecting changes both in the methods available for the infliction of widespread harm and those things that we value as contributing to valuable human lives. Much of the focus in the discussion around cyber conflict has centred on the new methods that are made available by information and communication technologies (ICTs) and our increasing reliance on those technologies. Less attention has been directed to the way our interests are being altered by the digital revolution. Increasingly, informational interests or cyber interests play a significant role in defining the nature of our lives. Increasingly, our welfare and wellbeing are determined by informational concerns. The digital revolution is changing both the ways in which we can fight and the things over which we fight. A successful ethical theory in this arena must accommodate both cyber methods and cyber interests.

The question that will be investigated is whether current ethical theories that concern themselves with war and large-scale conflict are able to accommodate the transitional changes that are implicit in the digital revolution. How do the changes that have been introduced by cyber conflict affect our ethical assessments? Will the continuing changes challenge even further the relevance of the ethical theories that have previously been used?

The analysis is in two parts. The first part constructs a framework, the cyber interests framework, that allows cyber conflict to be ethically assessed while considering both cyber methods and interests. The starting point of Part One is that cyber conflict can create widespread harm to groups and populations, and this justifies an ethical analysis. Frameworks that reduce or abstract harm to a single metric such as killings do not fully represent the harm that such conflict can inflict. In particular, they cannot adequately represent the harm of cyber conflict. Only a pluralistic and multivariate assessment of harm can reflect the diversity of our significant interests. Those interests must include families of cyber interests such as personal data control and public data access. While such a framework is not as clinical as a monist assessment of harm, it more adequately reflects those things that contribute to providing flourishing human lives and therefore more adequately reflects the multivariate ways in which harm can be inflicted.

Part Two investigates the effect that this cyber interest framework has on existing moderate theories of conflict. Three areas of concern that are essential to a successful theory of conflict are analysed using the understanding created by the cyber interests framework. Those areas are a prescription against excess (which is represented formally by the principle of proportionality), a requirement that only those liable are harmed (which is represented in the principle of discrimination), and various

demands that the infliction of harm is carried out for good reason (which are grouped under the heading of the principle of just cause). It is argued that any plausible theory of conflict must have some implementation of these principles and a theory or group of theories that is unable to implement them in a principled way cannot be successful in this context. This is the reason for the primary conclusion of the analysis: moderate theories of conflict cannot be applied to cyber conflict in a principled manner and the challenges that prevent the application are inherent rather than transitory.

In order to reach this conclusion, certain premises are assumed. Although these premises form the landscape of the argument presented, they are not directly defended. I believe that they are uncontentious, but, of course, there may be people who disagree. Those people are going to find the following chapters heavy going because the premises might be regarded as threads that reappear throughout the argument and, to a certain extent, provide a foundation for the argument. I will call them the axioms of the project.

We value flourishing human lives

The basic focus for concern is flourishing human lives. Whatever moral framework is applied, the underlying good is a flourishing human life. The value of flourishing lives may be expressed in terms of welfare and well-being. It may be expressed in terms of personal human interests or human rights. It may be expressed in terms of capabilities and functionings, or in various other ways. The underlying premise, that we value flourishing human lives, is agnostic of any specific ethical theory. It echoes the lay thought that we value having our lives go well although the breadth of that thought allows various interpretations.

Harm as a degradation of flourishing

Harm is a multivariate and complex concept. In what follows it is regarded as a degradation in the basic moral quanta, flourishing human lives. That definition of harm can be equally appropriate on the individual level or the level of groups or populations. This concept stems from the intuitive belief that if an increase in flourishing is seen as a benefit, then a decrease in flourishing must constitute harm.

We fight about those things we value

This premise seems intuitive. It is particularly relevant in assessing those things that might be regarded as the social or political outcomes that are the aims of conflict. If we are to argue that certain harms may be justified by certain outcomes - which is the approach of moderate theories of war - then those outcomes are assessed in terms of those things that we value. The twist that will be put on this understanding is that, if those things that we value change, then the things that we are willing to fight over will change.

We are in the midst of a digital revolution

The digital revolution is taken to be the societal changes engendered by the use of computers and ICT technologies. Personal computers, the Internet, mobile phones, smart systems, big data storage and analysis are all constituent elements of the digital revolution. The aim is not to engage thoroughly with the analysis of the digital revolution in terms of empirical analysis of the changes. Nor is it to engage in deep sociological analysis of the changes that are being wrought overall. Nor is it to engage with metaphysical analysis of the effects of the change on what it is to be human. The aim here is not to provide a substantive overall analysis of this digital revolution. The prime conclusion that is most important from this concept is that society is changing in significant ways and that as society changes the things that we value change. Also, worth emphasising at this early stage is that we have not yet experienced the entirety of the digital revolution. There have been major societal changes, but these changes are continuing, and the revolution is likely only to have just begun.

Finally, this is an undertaking in applied political philosophy. As discussed later, this provides a particular lens through which the issue is viewed. It is worth mentioning from the start, however, that it is not a work of legal theorising, but does allow that our legal systems and judgements elucidate the issue. Nor is it a work of sociology, but again allows that sociology is an appropriate methodology for discussing many of the issues, in particular the societal changes of the digital revolution. Likewise, it is not a purely empirical study of the issue, as this is left to others. It is political because the type of large-scale conflict with which it is concerned is inevitably political in the broadest interpretation of that word. It is philosophical because it attempts a conceptual analysis that is informed by the pragmatic realities of cyber conflict. It engages with an existing framework, Just War Theory, and uses the realities of large-scale cyber conflict to test that framework in a modern context.

Part One

1. Just War Theory and Cyber Conflict

An aggressive war is the great crime against everything good in the world. A defensive war, which must necessarily turn to aggressive at the earliest moment, is the necessary great counter-crime. But never think that war, no matter how necessary, nor how justified, is not a crime. Ask the infantry and ask the dead.

Ernest Hemingway (1946)

1.1. The Question

The digital revolution continues to cause significant societal changes. Informational assets have become increasingly important in our lives. These assets include both the data which influences our lives and also the systems which we use to manage that data. There have also been significant changes in the range of methods by which these assets can be influenced. Digitally stored data and systems can be influenced in ways that were never previously possible. Understanding the changes that are occurring involves consideration both of the changing informational interests in our lives and the changing methods by which those interests can be influenced.

This project concerns the way that those changes affect our existing moral frameworks of large-scale conflict. The type of conflict that is intended is that which imposes harm. Large-scale conflict is the type of conflict that imposes harm on groups or populations rather than just on the individual. It is this widespread imposition of harm that defines the scope and justifies the application of ethical analysis.

A wide definition of cyber conflict is used that is based on a specific understanding of the word 'cyber'. As will be defended later, cyber is taken to refer to the relationship between humans and information. A vague definition in which it is assumed that cyber refers to 'computer stuff' is inadequate and misleading. More accurate is a definition that claims that cyber refers to information communications technologies (ICTs). Current ICTs are the prevalent interface between humans and information, and the definition is narrow because of its reliance on current technologies. A more general, conceptual definition that is based on the relationship between humans and information and does not limit itself to particular technologies is more valuable in this fast-changing environment. Cyber interests are defined as informational interests that are relevant to human lives. Cyber methods are methods that target the relationship of humans and data. Cyber conflict is therefore any conflict that either targets the relationship between individuals and data, or uses that relationship.

The definition of cyber conflict outlined above is broad. It does not exclude actions that have either physical targets or physical methods. An attack on a dam using cyber methods qualifies. Conversely, a physical attack on informational assets also qualifies. An example might be the physical destruction of servers storing significant information. The defining characteristic of cyber conflict is that either cyber methods or cyber interests (or both) play a key role in the scenario.

This wide definition does not exclude what might be regarded as a paradigmatic example of cyberwar. Take, for example, a case in which the control system for a missile control system is subverted by digital techniques. The control system is interpreted as a cyber interest and the digital techniques as cyber methods. This example conforms to the wide definition. However, the definition also allows that there are other forms of cyber conflict than this militaristic example. Any action or campaign that creates widespread harm and involves cyber interests or cyber methods is taken as an instance of large-scale cyber conflict.[1] Cyber conflict is not only the domain of military actions.

Cyber conflict is developing rapidly, its significance is growing, and it raises its own distinct questions. Can our existing theories of war and large-scale conflict accommodate cyber conflict? Or is there something inherently different about cyber conflict that makes these theories inapplicable or inappropriate in the new context? Given the changes that are resulting from the digital revolution, is it reasonable to continue to use existing ethical frameworks to understand and mediate our large-scale conflicts? In response to these questions, this project argues that the concepts that constitute cyber conflict challenge the foundations of current ethical theories of conflict and make their use problematic or inappropriate in the arena of cyber conflict.

A paradigmatic example of our current ethical frameworks is provided by Just War Theory (which for simplicity will be termed JWT). JWT is a specific type of theory that attempts to balance the moral costs of war with the moral benefits. It claims that war may be morally permissible in circumstances in which the benefits are great enough to justify the costs. In this way, theories of this type are not extreme, and are termed moderate theories of war in this project. The analysis in the following chapters is

[1] The reasons for the framing of the scope in terms of large-scale conflict rather than in terms of war will become increasingly clear in this chapter but are centered around the idea that some large-scale conflicts may be hard to designate as war in the context of modern geopolitical rivalries, but nonetheless involve the imposition of significant harm. Large-scale conflict is a broader concept than the somewhat narrow conventional definitions of war. For simplicity, the terms are used as interchangeable in a broad sense, and only when the distinction is required is it strictly enforced. Often ‘war’ will be used to simplify the language.

aimed specifically at JWT but may equally be applied to other moderate theories that balance harms and benefits.[2]

Ethical theories will define how actions in large-scale context might be morally assessed. It is assumed that a requirement for a successful ethical theory is that it provides an assessment of whether engagement in large-scale conflict is permissible and how we should behave once engaged in this type of conflict. An ethical theory is appropriate because there are significant moral costs associated with war. Those costs include the mass horrors with which we are familiar from history but also include less visible, yet pernicious, harm across populations including economic, psychological, and emotional harms. The need for an ethical theory does not exclude other forms of theory, for example physical, economic, sociological, etc. Ethical concerns are only one way that war and conflict can be analysed, but they are the way that is chosen in this project.

The overall path of this project is to define and analyse certain essential elements of any plausible philosophical moderate ethical theory of large-scale conflict. These elements are based on three pre-theoretic thoughts or beliefs. The first is a prescription against excessive actions. For instance, all other things being equal, the threat of imprisonment of a single citizen would not justify launching a nuclear strike on a major city. In more technical terms, we will see that this thought grounds the principle of proportionality. The second is a demand that harm should only be inflicted on appropriate targets. Phrased differently, this thought is that harm should not be indiscriminate, and we will see that it grounds the principle of discrimination. The third is that there must be a good reason to inflict harm. The harm cannot be without meaning or purpose. We will see that this thought can be difficult to define in more precise terms. For the present purposes, this thought is corralled under the heading of the principle of just cause. None of this excludes the possibility that other concerns may also be relevant, but it will be argued that without an implementation of any one of these pre-theoretic thoughts, a moderate theory cannot be either plausible or successful.

In the context of cyber conflict, this is unfortunate because implementations of proportionality and discrimination are inappropriate. They are inappropriate because those principles either cannot be applied in a principled manner or return results that are indeterminate. This is due to the inherent nature of cyber methods and cyber interests. It is not a transitory issue or a mere curiosity. There are fundamental characteristics of cyber conflict that cause these failures. The result is that Just War Theory and other moderate theories cannot be applied to cyber conflict.

---

[2] These are moral costs and benefits. Because of this fact moderate theories are not limited to consequentialist theories. Even a purely deontological theory might still attempt to balance moral costs.

1.1.1. So What?

How we might react to this failure of moderate theories to accommodate cyber conflict is dependent on the way that those theories are conceptualised. Specifically, it is important whether a theory is a theory of allowance or limitation. A theory of allowance emphasises the moral cost of war but allows that this moral cost can be justified in certain circumstances. A theory of limitation accepts that there may be morally significant benefits in engaging in war but limits that engagement to cases in which the costs are not so high that they overwhelm the justification. Those of realist leanings are likely to conceptualise moderate theories as theories of limitation while those of pacifist leaning are likely to conceptualise them as theories of allowance.

It is possible that a theory of allowance and a theory of limitation look very similar. In fact, it may be that they both result in the same pragmatic prescriptions. What is argued in the following paragraphs is that the difference becomes apparent when the theories fail, either through inaccuracy or indeterminacy. Both allowance and limitation rely on assessing the criteria on which a judgement might be based. In cases where that assessment is flawed, the 'fallback' position will be significantly different.

In the West, certainly before the twentieth century, theories of allowance are predominantly based in the Catholic tradition. This is a tradition of war theory that starts with Augustine and includes notable theorists such as Aquinas, Vitoria, and Suarez,[3] and attempts to answer the challenging questions for Christians that war raises. The concept that underpins this tradition is that peace is desirable, and war is only allowed in very specific and carefully defined cases. The current catechism (Catholic Church 2019:2302-2317) in a section entitled "Safeguarding Peace" starts by reminding readers that killing is a mortal sin. It is clear that the tradition is one of the allowance of harm only when the benefits of engagement can justify it. This is not to say that only Catholic theories are, or can be, theories of allowance. Many secular theorists follow this line which is made plausible by the fact that the moral costs of war are extreme. A.J. Coates (2016:21) explains, "However ‘just,’ no war is ever so pure or ever so untainted as to be entered into without grave misgivings." For many, this is the most salient fact about war, and it demands that widespread harm can only be permissible when tightly defined criteria are met.

In contrast, a theory of limitation states that we are empowered to pursue our aims and that only in certain cases should this pursuit be limited by other concerns. The concept that usually underpins these theories is that one has a right to preserve one’s

---

[3] See, for example, *The City of God* (Augustine 2014), *Letters* (Augustine 2003), *Summa Theologica* (Aquinas 2006), *Suárez on just war* (Reichberg 2011), *Opera omnia* (Suarez 1958), *Francisco de Vitoria on the “Just War”* (Mantovani 2017), *De Indis Et de Iure Belli* Relectiones (Vitoria 1964).

person and one's interests. Hobbes[4], in particular, took this reasoning to an extreme level in which most actions could be thought of as preservation of one's interests. Only extreme cases limit this freedom to act according to one's aims.

Theories of allowance and theories of limitation are both moderate theories. Although a theory of allowance may start from a seemingly pacifist stance, it is not strictly a form of pacifism in that it conforms to the idea that harm can be justified by the value of the outcome in certain circumstances. Likewise, theories of limitation are not realism in the strictest sense as they accept the idea that the inevitable harm of actions can make the pursuit of one's aims impermissible. Both are moderate in that they aim to compare the value of the consequences of actions against the harm of those actions. They use this comparison to define circumstances in which harm may be either allowed or limited.

Perhaps there is no difference between a theory that limits harm in certain circumstances to one that allows harm in certain circumstances? Is there a difference between saying, for instance, that an action is permissible if, and only if, certain criteria are met, to saying that an action is impermissible if, and only if, certain criteria are not met? The answer is that they are not the same claim. In logical terms this is seen by the fact that if A then B does not imply if !A then !B. In common terms, 'stop if the light is red' is not the same as 'do not stop if the light is not red.'

For moderate theories of war, the principles proportionality, discrimination and just cause provide some of the criteria that are used for determining either limitation or allowance. For example, a limiting theory claims that an action is not permitted if it is indiscriminate. An allowing theory claims that an action should be allowed only if it is discriminate. Imagine for a moment a situation in which the principle of discrimination cannot be applied or returns indeterminate results. Our responses in such a situation - in which we are not getting adequate guidance from the principle - are likely to be dependent on whether we understand the moderate theory as an allowance or a limitation. In the case of an allowing theory, the allowance will not be justified, and the action is impermissible. In the case of a limiting theory, the limitation will not be justified, and the action is permissible.

A persuasive argument will be presented in the following chapters that the application of the principle of discrimination (Chapter 4) and the principle of proportionality (Chapter 5) is inappropriate in the context of cyber conflict. At the very least, its application creates indeterminate results. This may leave one wondering if one is forced towards a more fundamental form of realism, or if one is forced towards a more fundamental pacifism.

It is outside the scope here to resolve the overall tension between realism and pacifism and any hope for that resolution is naive. It may be impossible to adjudicate

---

[4] See, for example, Leviathan (Hobbes 1968) and *Philosophicall rudiments concerning government and society* (Hobbes 1651).

in any formal manner the relative values of the freedom to pursue one's own interests and the imperative not to create harm to others. However, the analysis and the frameworks that this project develops do have something to say about this issue. While it is not the main thrust of this project, Chapter 6 explores how we might view the failure of moderate theories in the light of these issues. It does so by outlining a certain set of challenges that emerge for those that feel that the correct course is a more fundamental form of realism.

The fact that Chapter 6 is exploratory does not diminish the impact of the project. The primary aim of the project is to undermine the use of moderate theories in the context of cyber conflict by showing that their application is either inappropriate or results in indeterminacy. The conclusion is that the principles of proportionality and discrimination are not applicable in the context of cyber conflict and that moderate theories are therefore inapplicable in this context. A subsidiary aim is to present the challenges to the realist route with the aim of making the pacifist route comparatively attractive.

A critic of the method of this project might initially claim that inappropriate weight is being placed on three pre-theoretic thoughts or on the three principles that they ground. The critic might say that even if these principles do not work the way that they are intended in the context of cyber conflict, there are other principles that might replace them. Another line that the critic might take is to claim that these principles are not needed, and ethical judgements can be made without them.

Both these challenges are answered by the fact that the pre-theoretic thoughts are intuitive and plausible. The starting point is the pre-theoretic thoughts rather than the specific principle that they might ground. There is widespread discussion and disagreement about the principles once they are formulated in technical terms, but the plausibility of the underlying thoughts is rarely, if ever, doubted. It may be challenging to define what an excessive action is, but an ethical theory that does not limit excessive actions - actions that imposed significant harm beyond any justification - is untenable. Likewise, although the correct available targets are often difficult to demarcate, a theory that does not justify the imposition of harm to the relevant individuals or groups - one that allowed indiscriminate harm - is untenable. A good reason for the imposition of harm is also hard to define, but a theory that allows the imposition of harm with no good reason is untenable.

Untenable in this context references the fact that those theories are intended as ethical theories. There are standpoints such as the non-philosophical type of realism that denies the relevance of morality in this context, that might allow one to discard one, or all, of these thoughts, but that type of standpoint is outside the scope of an ethical philosophical analysis of the issues. It is taken that a successful ethical theory must accommodate our most simple pre-theoretic thoughts. The three thoughts that have been outlined are sufficiently persuasive that they demand to be implemented in a successful theory. A theory that does not implement them, for instance one that allows excessive and indiscriminate harm, cannot be plausible. This is a high-level analysis that is independent of the particular theory.

The most prevalent moderate theory of war has been JWT and the following section how it has traditionally implemented and instantiated these three thoughts.

1.2. Just War Theory

> **Even in waging war, cherish the spirit of peace-maker; that, by conquering those whom you attack, you may lead them back to the advantages of peace.**
>
> **Augustine (Letter 189: A.D. 418)**

1.2.1. Early Historical Context of JWT in the West

The history of just war theory in the West is extensive and the various forms that the theory has taken have led some commentators to believe that it is better to talk in terms of the just war tradition.[5] Either way, JWT is an acronym that will be used in this project, covers both bases, and allows the distinction between tradition and theory to be sidestepped. The documented history of JWT in the West is often described as having a foundation in Catholicism that starts with Augustine.[6] Augustine certainly is a major influence even though he built on an existing body of thought and, as Mattox (2009:161) claims, he "does not present his views on just war as a unified theory".

Augustine is described by Christian Tornau as a "Christian Neoplatonist, North African Bishop, Doctor of the Roman Catholic Church" and continues immediately to describe him as "one of the towering figures of medieval philosophy whose authority and thought came to exert a pervasive and enduring influence well into the modern period,

---

[5] For example, David Rodin says, "But the Just War Theory is, in reality, many theories. It is more accurate to talk of the 'just war tradition' rather than the 'Just War Theory', for it includes a large number of diverse yet related positions stretching from the theological writings of Augustine and Aquinas, via the legal treatises of Grotius and his contemporaries, to the modern secular account found in writers such as Michael Walzer." (Rodin 2002:103-104)

[6] There is evidence that the tradition is older than Augustine - still, one needs to start a story somewhere and, in the West, a possible starting point is with Augustine. Larry May (2018:1) opens the introduction to The Cambridge Guide of the Just War with the claim that "The idea of a 'just war' has been with us since Mesopotamian times." Christian Torhau (2020: §8) in the Stanford Encyclopedia of Philosophy describes Augustine's work in this context as a "Christian reinterpretation of the traditional Roman Just War Theory".

and even up to the present day" (Tornau 2020). Augustine's influence on JWT was, and continues to be, extensive. The word 'pervasive' is an accurate description. In fact, the first section of John Mark Maddox's chapter (Maddox 2018) in the Cambridge Guide to the Just War is titled "Augustine's Pervasive Influence." Given that over 1500 years have passed since Augustine wrote there might appear to be some question as to whether his views are relevant in the current world. However, it is the pervasive and enduring nature of Augustine's thoughts on JWT that makes him noteworthy.

This emphasis on Augustine does not diminish the philosophers that came after him in this Christian tradition. Mattox (Mattox 2018:28-31) describes the extensive influence of Augustine on thinkers in "Late antiquity and the Middle Ages" including Gratian and Aquinas. Many people who are not philosophers and who have a Catholic upbringing regard JWT as only concerning Augustine and Aquinas and are unaware of later philosophical developments. However, with the peace of Westphalia in 1648, the background to war changed and new thinkers were less compelled than Augustine to derive their theories of just war from divine principles. Theorising about JWT became increasingly secular - a process that resulted in the modern emphasis on legal interpretations based on international law. Despite this move towards secularism, the influence of Augustine and the early Catholic tradition cannot be overstated. As we will see, although the justifications for the elements of the theory may have changed, the core conceptual elements of the theories have remained remarkably constant.

It is worth an aside regarding current Catholic catechism (Catholic Church 2019). The relevant section starts with Article 2302 which states, "By recalling the commandment, 'You shall not kill,' our Lord asked for peace of heart and denounced murderous anger and hatred as immoral." It is significant that the discussion of peace and war starts with this statement. The section continues with the articles most relevant to JWT, in particular, Article 2309:

> The strict conditions for legitimate defense by military force require rigorous consideration. The gravity of such a decision makes it subject to rigorous conditions of moral legitimacy. At one and the same time:
>
> - the damage inflicted by the aggressor on the nation or community of nations must be lasting, grave, and certain;
>
> - all other means of putting an end to it must have been shown to be impractical or ineffective;
>
> - there must be serious prospects of success;
>
> - the use of arms must not produce evils and disorders graver than the evil to be eliminated. The power of modern means of destruction weighs very heavily in evaluating this condition.

As will be seen in the following subsection that concerns itself with the structure of JWT, there are salient similarities between Augustine's formulation of JWT and the current Catholic catechism. It remains noteworthy that the Catholic interpretation of JWT has always been an allowance of harm only in cases where that harm has been justified.

As an aside to this aside, Pope Francis, in 2020, signed his third encyclical entitled *Fratelli Tutti* (Pope Francis 2020). It is widely perceived that in *Fratelli Tutti,* Pope Francis distances the Catholic Church from the doctrine of JWT. Theologian Terrence Rynne is quoted (McElwee 2020) as claiming that "He throws it, to all intents and purposes, in the ash can." It is unclear what doctrine will replace JWT in Catholic teaching but it is clear that the issue is in transition. It is interesting to observe that the Catholic Church seems to be faced with the choice between pacifism and realism that was described previously. The obstacles that face the adoption of a philosophically based form of realism appear overwhelming for the Catholic Church. This echoes the underlying premise of this project in that, if moderate theories prove untenable, there remains a choice between realism and pacifism, and there are strong challenges to those aiming to choose a form of realism.

The purpose of these asides has been to demonstrate that the nature of the Catholic influence has not changed significantly since Augustine - although it might be changing in the 21st century. In the next subsection, we will see that Augustine's formulation of JWT is also remarkably similar to modern secular formulations. The structure, the language and many of the actual criteria of assessment overlap significantly.

There might be different explanations for the overlap between Augustine's JWT and modern versions. Perhaps Augustine was brilliant, and his formulation is so persuasive that it influenced all subsequent thinkers. Another thought might be that there are certain elements that are essential for any plausible interpretation of JWT. Of course, both may be true to some extent, but it is the second idea, that there are essential elements of a moderate theory that reinforces the claims of Section 1.1.1 - that some implementation of proportionality, discrimination and just cause in their broadest interpretations is essential for a successful moderate theory.

There are possible concerns about the way that JWT theory has developed in the West. Aspects that are particularly noteworthy are the use of the phrase 'just war,' a binary division between just and unjust sides in a conflict, and a strong division between reasons to go to war and behaviour once engaged in war.

In its modern interpretation 'just' can be interpreted as righteous, equitable, fair, justifiable, lawful, and even perfect (Harper n.d.). In the case of Just War Theory the exact meaning of 'just' is rarely discussed. Certainly, the implications of, and emotional response to, a 'righteous war theory' might be different to those of a 'lawful war theory,' or a 'perfect war theory.' One's emotional response to the idea of JWT may be governed by how one interprets the word 'just.' Despite this ambiguity, the use of the phrase 'just war' continues.

Even in modern and sophisticated literature, it is common to find a binary division between the just and unjust sides of a conflict. For instance, McMahan (McMahan 2015:4) frequently uses phrases such as “Many civilians on the unjust side in a war”. The language used in much of the discussion remains based on an archaic and binary division between just and unjust that does not reflect the reflective truth that the division is not as clearly bounded as that language might suggest. Often, neither side in a large-scale conflict is absolutely just nor absolutely unjust. It is important to note that the language that is used is a powerful rudder to the conversation and that this binary division has the tendency of simplifying complex situations to the extent that it is possible that the results are due to the simplification rather than the situation itself. The terminology is not neutral in any discourse.

Another terminological issue that is pervasive is the distinction between Jus ad Bellum and Jus in Bello. Jus ad Bellum refers to justice in entering into war. It concerns the question ‘are we morally justified in entering into war?’ Jus in Bello refers to justice during war. It concerns questions as to how we should behave in a war. JWT concerns itself both with how wars are fought and why wars are fought. In broad brush strokes, the justification of ‘how’ is represented by Jus in Bello while the justification of ‘why’ is represented by Jus ad Bellum. The use of Jus ad Bellum and Jus in Bello is nearly universal in Western literature. Using these terms is clear for academics and experts in international law. It is less clear for the majority of the population. It seems to me that our theories of war and other conflict should be easily comprehended by the majority of the population and not a minority or elite.

Up to this point, the persistent concerns regarding the development of JWT have centred around the terminology used to describe those persistent concepts. The use of the phrase ‘just war,’ the binary division between the just side and the unjust side of a conflict and the use of the phrases Jus ad Bellum and Jus in Bello all have worries associated with them despite the plausibility of the ideas that they represent. These worries cannot be laid entirely at Augustine’s feet,[7] but there are legitimate worries concerning the language that is used to describe the ethical consideration of large-scale conflict. One needs to be wary of the influence that this language has on the discussion as a whole.

Because of the worries about the language and formulation of particular instances of JWT this project will be concerned with a more general categorisation of theories as moderate theories of large-scale conflict. JWT is certainly a member of that category. Whether all moderate theories may be called JWT is a question that is left to others to answer. It is a possible approach to say that all moderate theories are theories that

[7] Partly because the development has continued for over 1500 years since Augustine and partly because closer inspection reveals that Augustine did not use the word 'just' and that the Greek 'iusticia' has different cultural connotations, that he did not use the terms Jus in Bello and Jus ad Bellum, and that he was not as firmly committed an idea of non-relative justice as a binary division between just and unjust implies.

aim at the justification of war or large-scale conflict and therefore qualify as instances of JWT. It might also be possible to argue that just war theories are a subset of moderate war theories. This is largely a semantic discussion, and it can be avoided by talking in terms of moderate theories of large-scale conflict. This has the added advantage of removing the emotive content of some of the language associated with JWT.

War and large-scale conflict present a moral problem. They inevitably involve the imposition of harm on populations, and yet at times, a desired outcome is only attainable through large-scale conflict. In fact, for those uninterested in ethical concerns, a successful war might be described as the attainment of one's aims by means of the imposition of harm on others. This leads to a higher-level historical worry about JWT. While the idea of a justified war makes conceptual sense, there is a fear that the purpose of JWT is to permit war rather than limit it; that although JWT is often justified as a theory of limitation, it is implemented as a theory of allowance, that all these clever and complex words are simply a method of permitting the strong to do what they will. Can this fear be dismissed as unfounded? Well, it certainly could if, historically, all the wars that had been declared 'just' were interpreted later as objectively justified and legitimate. It certainly could if the delineation of what is defined as a 'just war' was sufficiently clearly defined and agreed upon that there was no dissension in its assessment. If neither of those concerns, or their like, are met, then the fear remains legitimate. We must be cautious that the idea of 'just war' is not simply an elaborate technique for allowing the strong to coerce the weak under cover of a supposedly moral framework. For this reason alone, an ethical analysis of the ability of existing theories to accommodate developments in society is warranted.

With this in mind, it is worth noting that there are elements of Augustine's thoughts that have not been entirely persistent or persuasive. Augustine's starting point was one of pacifism rather than realism.

> But, say they, the wise man will wage just wars. As if he would not all the rather lament the necessity of just wars, if he remembers that he is a man; for if they were not just he would not wage them, and would therefore be delivered from all wars. For it is the wrong-doing of the opposing party which compels the wise man to wage just wars; and this wrong-doing, even though it gave rise to no war, would still be a matter of grief to man because it is man's wrong-doing. Let every one, then, who thinks with pain on all these great evils, so horrible, so ruthless, acknowledge that this is misery. (Augustine 2014:311)

It is certainly worth remembering that we should lament the necessity of even supposedly just wars. This simply reflects the epigraph in which Hemmingway says, "But never think that war, no matter how necessary, nor how justified, is not a crime." With the secularisation of JWT, this theme has not been consistently persistent.

As has been outlined above there are legitimate concerns about the purpose, language, and historical success of JWT. To understand how these concerns may play out in the context of cyber conflict a certain amount of groundwork is necessary.

The next section starts that groundwork by providing a summary of the traditional structure of JWT. If JWT is to be seen as successfully delineating just wars, it is this structure that will be critical.

1.2.2. Structure of Just War Theory

Augustine did not formulate his thoughts regarding JWT into a single treatise and his version of JWT has been reconstituted by later commentators. Nevertheless, it is widely taken that he suggested certain criteria that were essential for both the entrance into war and one's behaviour during war (e.g., Mattox 2009 and Mattox 2011). Although he did not use the terms himself, his criteria are usually divided into criteria that determine the justice of entering into war (Jus ad Bellum) and criteria that determine the behaviour of participants in war (Jus in Bello). Augustine's principles of Jus ad Bellum are:

- A Just Cause
- A rightly intended will
- A declaration of war by a competent authority

And his principles of Jus in Bello are:

- Proportionality
- Discrimination of proper objects of violence (e.g., combatants not civilians)
- Good Faith (e.g., honesty and a lack of treachery)

A full description of Augustine's principles is left to others. At no point in the history of Western JWT has there been an uncontested formulation and there remains a good deal of discussion about how any list of perhaps necessary criteria might be defined or delineated, or even if a list of criteria is appropriate. More recent guides to JWT, such as The Cambridge Handbook of the Just War, steer away from prescriptive lists of elements. The division between Jus ad Bellum and Jus in Bello is widely used and understood, but additional categories such as Jus ex Bello (justice while withdrawing from war. - see Blum and Luban 2015, Mollendorf 2008, Rodin 2015 & Statman 2015) and Jus ad Vim (justice in measures short of war - see Galliott 2019) are appearing in contemporary renderings of JWT.

With that in mind, the following structure of JWT is offered as an overview of the salient aspects of conventional JWT rather than as a definitive listing. It is an attempt to provide a general statement of a standard view and is at least partially based on the article in the Stanford Encyclopedia of Philosophy by Seth Lazar (Lazar 2020) who does his best to provide a neutral description.

Jus ad Bellum

- Just Cause
- Competent Authority

- Right Intention
- Probability of Success
- Last Resort
- Proportionality

Jus in Bello

- Necessity
- Proportionality
- Distinction/Discrimination

As I have said, the inclusion or omission of particular principles is contested. Likewise, the precise definition and interpretation of all the principles are contested. With that in mind, and because proportionality, distinction and just cause are dealt with in much greater detail later, the following outlines are given only as a map of the landscape.

‘Just cause’ is perhaps one of the most deceptively difficult of the group to pin down. In pre-theoretic terms it is easy - just cause simply means one has a worthy reason to enter into war. A more technical definition is difficult. Lazar (2020 Section 2.5) uses the phrase, "the war is an attempt to avert the right kind of injury." This captures the idea that a just cause must meet certain criteria but does not specify what those criteria are. It is the specification of those criteria that hides the difficulties in just cause. Just cause is one of the focuses of Chapter 6.

The idea of ‘competent authority’ replaces the idea of a declaration of war. This might be regarded as a lowering of the bar but is simply a response to the realities of modern conflict. Competent authority demands that the people or government initiating the engagement in war have the authority to do so. For instance, an opposition party in a conventional democracy would not have authority. The idea of competent authority is unproblematic in a simple world-conception based on Westphalian nation states but is significantly more complex in a world of civil wars, terrorist organisations, freedom fighters and cross-border organisations and groups.

Lazar sketches ‘right intention’ with the phrase, "that entity intends to achieve the just cause, rather than using it as an excuse to achieve some wrongful end." Broadly, this is how right intention is understood in most contemporary writing. I suspect that Augustine’s idea of a ‘rightly intended will’ was somewhat different and carried what might now be considered moralistic and religious overtones. In the modern conception, right intention is a limiting principle that aims to prevent actors using legitimate justifications for war to pursue war for other reasons. One route to challenge this would be to promote the idea that intent is not relevant in the moral assessment of an action. There are also potentially deep epistemic problems in assessing the true reasons of those making the decision to engage in war.

‘Probability of success’ simply states that the intended war is sufficiently likely to attain its ends. The principle aims to prevent engagement in wars that are unlikely to

succeed but will inevitably impose significant harms. Many commentators believe that this principle can be rolled into proportionality. As Thomas Hurka says,

> The proportionality conditions are actually more important than this initial account suggests, since, if formulated properly, they can incorporate the other just war conditions about consequences. Imagine that a war has no chance of achieving any relevant goods. This fact, which makes it violate the reasonable hope of success condition, surely also makes it disproportionate, since its destructiveness now serves no purpose whatever. The same is true if the war has only some small probability of achieving relevant goods, since then its expected harm is excessive compared to its expected good. If it takes account of probabilities in this way, as on any plausible view it must, the ad bellum proportionality condition incorporates hope-of-success considerations, and it can also incorporate last-resort considerations. (Hurka 2005:37)

'Last Resort' is exactly that. Lazar phrases it as '"there is no other less harmful way to achieve the just cause." As seen above, it seems plausible that last resort can be incorporated into proportionality.

'Proportionality' is a limiting principle that aims to disallow excessive actions. It appears both in the Jus ad Bellum criteria and the Jus in Bello criteria. It balances the benefits of the outcome against the harm produced by the action. Lazar's phrasing hints at the technical challenges in defining proportionality, "the morally weighted goods achieved by the war outweigh the morally weighted bads that it will cause." As proportionality is fully discussed in Chapter 5, further discussion is left until then.

'Necessity' is a Jus in Bello principle that demands that there is not another less harmful course of action to arrive at the same goal. It compares the 'morally weighted bads' of the various options and demands that the option with the least is pursued.

'Distinction' or 'discrimination' is a principle that demands that only the correct individuals are intentionally targeted. Historically, this has meant that only military personnel, or 'combatants', may be intentionally targeted, not 'civilians'. In modern discourse, there are problems raised in determining who exactly qualifies for immunity from attack, what the grounds for immunity actually are, and how to assess the criteria of intention. Discrimination is discussed more fully in Chapter 4.

As I write these outlines, it is apparent that even the broad brushstrokes used may cause some contention. The aim is to provide an overall map of the landscape rather than support any particular interpretation. Lazar, I assume, has the same aim in his outline, and yet there are areas that he leaves open to challenge. What for instance, does 'sufficiently likely' mean in his outline of probability of success? At this stage, I have simply tried to map the terrain with as much impartiality as possible and have aimed not to enter into, let alone resolve, any of the discussions of contentious issues.

The list presented is not intended to be exclusive or final. There are elements, such as a prohibition of methods that are inherently wrong (*malum in se*), that are often

included. The elements are often defined in somewhat different manners or divided in different ways. A list of elements is contentious.

What is relevant to this project is that the three pre-theoretic thoughts that we started with are always reflected in conventional JWT. The prohibition of excessive actions is instantiated in proportionality and necessity. Last resort and probability of success can also be conceptualised in terms of prohibition of excess. For instance, Thomas Hurka (Hurka 2005:37) argues that the criteria that involve consequences can all be incorporated into proportionality. A demand for correct targeting is evident in the principle of discrimination that at its most basic states that harm cannot be indiscriminate. The demand that the imposition of harm be carried out for a good reason appears in the principle of just cause and possibly right intention although these elements are perhaps the most contentious. All implementations of JWT address the pre-theoretic thoughts and only a small minority lack a direct implementation of them.[8]

1.2.3. Discussions and Divisions of Just War Theory

It is also worth mapping out the shape of the major divisions in the recent discourse about the ethics of war. The philosophical interest in JWT was re-ignited by the book, Just and Unjust Wars, by Michael Waltzer (Walzer 2006a - 1st Ed. 1977). The historical context is important. Waltzer was writing in the immediate aftermath of the Vietnam War and overtly claims that he started thinking of war as an activist rather than a philosopher.[9] His approach was based very much on historical and real examples and aimed to "set out the moral argument about war in a quiet and reflective way." (Walzer 2006a:xx) The often-neglected full title of the book is "Just and Unjust Wars: A Moral Argument with Historical Illustrations". The nature of Walzer's book and the responses to it have shaped the subsequent discussion. The standpoint adopted by Walzer has been termed 'traditionalist'.

Walzer had laid out positions on key issues such as national defence, humanitarian intervention, discrimination, and combatant equality. National defence was accepted as a core right of the nation state.[10] Humanitarian violations could justify

---

[8] For instance, Uri Steinhoff has argued that Just Cause amounts to little other than meeting the other requirements of JWT. See. for example, Steinhoff 2014

[9] The first lines of the original preface to *Just and Unjust Wars* reads, 'I did not begin by thinking about war in general, but about particular wars, above all about the American intervention in Vietnam. Nor did I begin as a philosopher, but as a political activist and a partisan.'

[10] "But most states do stand guard over the community of their citizens, at least to some degree: that is why we assume the justice of their defensive wars." (Walzer 2006a:54) Walzer defends the right to defence throughout *Just and Unjust Wars.*

intervention.[11] Discrimination between civilian and combatant was a central element of the morality of war.[12] Combatants were morally equal when they confronted each other regardless of the justness of their war.

All these traditionalist standpoints were rapidly challenged by what has become known as the 'revisionist' approach. The underlying premise for these revisionist challenges was that the traditionalist position did not supply a principled moral defence for key aspects of the traditionalist view. There was not sufficient ethical justification, the revisionists claimed, for some of the positions taken on key issues. Given that fact, it was argued, those aspects of the view had to be reassessed.

This re-assessment took two forms. The revisionists sought to challenge the conclusions of the traditionalist stance,[13] while others sought to provide the principled moral defence of traditionalist conclusions that the revisionists claimed was missing.[14] The positions adopted are so diverse that the categorisation of traditionalist and revisionist becomes increasingly challenging. However, the division between the traditionalist and revisionist standpoint emphasises that there are a number of ways in which we may think about war, and that there are axes along which our thoughts may be described.

The first axis of note is the distinction between whether we are talking about what the rules governing war should be, or whether we are talking about what the morality of war is. The first approach is 'institutionalist' in that its focus is the formation of morally guided institutions such as the law. The idea is that one creates those morally grounded institutions, and the individual is required to follow their guidance. The second approach is to think initially of the moral decisions and motivations of the individual or group and demand that actions are determined according to their moral reasons. The initial questions that the two approaches ask might be 'what should the rules for war be?' and 'what should we morally do in respect of war?'.

---

[11] "Against the enslave-ment or massacre of political opponents, national minorities, and religious sects, there may well be no help unless help comes from outside. And when a government turns savagely upon its own people, we must doubt the very existence of a political community to which the idea of self-determination might apply." (Walzer 2006a:101) Walzer devotes Chapter 6 of *Just and Unjust Wars* to interventions in which he attempts to define the cases in which such an intervention can be justified.

[12] In Chapter 9 of *Just and Unjust Wars,* Walzer analyses civilian immunity with the context of necessity. " Once the contribution has been plainly established, only 'military neces-sity' can determine whether the civilians involved are attacked or not. They ought not to be attacked if their activities can be stopped, or their products seized or destroyed, in some other way and without significant risk. The laws of war have regularly rec-ognized this obligation." (Walzer 2006a:146)

[13] Notable exponents of this approach are Rodin, McMahan, Luban, Fabre and others.

[14] Notable exponents of this approach are Lazar, Steinhoff, Benbaji and others.

In general terms, this project is not institutionalist. The questions it asks are considered in the scope of how we should think of war and large-scale conflict rather than what institutions are appropriate for the management or control of war.

The second axis is a methodological one. Walzer, as was noted above, based much of his discussion on historical facts. Put bluntly, historical facts are messy and do not always cleanly display the things that philosophers want to be displayed. Moreover, we import our own biases and interpretations to those examples. There is a motivation to use abstract and idealised examples rather than historical examples as these examples can be designed in an attempt to remove the kind of bias and lack of clarity that is associated with history. Those who support historical examples might argue that abstract examples are simply not relevant to the reality of war. Those who support abstract examples might argue that historical examples are inherently skewed by personal bias. It is worth noting that although the first revisionist challenges to Walzer were based on abstract examples, this mapping is not always the case. There are traditionalists that use abstract examples, notably Seth Lazar and revisionists who use historical examples, notably Cecile Fabre.

This project does not take a position in this discussion. This is partly due to the short history of cyber conflict and the fact that sufficient examples are not always available. Moreover, the value of abstract examples is evident in moral philosophising. However, where abstract examples are used, an attempt at assessing any conclusions in the context of real-world cases is also attempted. The danger of abstract examples is that they may become distanced from reality. This assessment of real-world cases after the use of abstract examples is an attempt to avoid this danger.

A third axis is defined by reductionism and exceptionalism. Exceptionalists, or non-reductivists, believe that the moral arguments that apply in war are different from those that apply outside war. War is an exception to ‘normal’ morality. Reductivists believe that the morality of war is grounded in the same moral factors as interpersonal morality.

A fourth axis is defined by individualism and collectivism. A collectivist will believe that an act may be associated with the collective rather than with the individuals that make up that collective. In opposition, an individualist will believe that an act is either reducible to the acts of the individual or that the significance of the collective is only assessed in terms of the individual. Collectivists usually treat the group as a moral entity in its own right whereas individualists do not.

This project treats war as an inherently political phenomenon and does not demand a reduction to personal morality. It is not denied that our rights and duties as groups may not be the same as those for individuals or reduce to our individual rights and duties. This does not necessarily imply that our individual rights and duties are not relevant to the discussion of large-scale conflict or war.

A similar path is taken between individualism and collectivism based on the idea that war is inherently political. The rights and duties of groups are significant in this context as are the rights and duties of individuals. Discussions as to which is the best or correct way of conceptualising large-scale conflict are left to others. If the arguments of this project are persuasive, they will be persuasive both to collectivists and individualists.

The four axes above might be viewed as intellectual schisms. This is somewhat unhelpful. After all, even a committed collectivist may find some value in viewing a particular problem from an individualist standpoint, even if only to analyse their collectivist intuitions more fully, and vice versa. For the most part, the axes represent a spectrum of ways of looking at particular problems. The 'perfect' conclusion would be palatable and persuasive to both collectivists and individualists for instance.

In terms of the nature of this project, some of this discussion is moot. The reason for that is that the target of the project is to analyse the effect of the digital revolution on particular principles of JWT. For instance, if one is able to show that cyber conflict makes the application of proportionality problematic, then that fact will largely be independent of one's position on these axes. An individualist is likely to have a slightly different interpretation of proportionality than a collectivist, but if it can be shown that proportionality on any plausible interpretation is inappropriate to conflict in the cyber domain then the difference between interpretations is largely irrelevant.

This project directly targets proportionality and discrimination with respect to cyber conflict. It is the tools of JWT - and other moderate theories of war - that are targeted rather than any particular interpretation of the moral implications of those tools. The aim is to demonstrate that two essential elements of any plausible interpretation of a moderate theory (a prescription against actions that are excessive and a demand that harm is only imposed on legitimate targets) are inapplicable in the context of cyber conflict. In order to achieve this aim, some groundwork regarding the context is required.

1.3. Cyber Conflict

1.3.1. Defining Cyber Conflict

As was explained in Section 1, a wide definition of cyber conflict is used in this project. Cyber conflict is any conflict that either targets the relationship between individuals and data, or uses that relationship. There are two aspects of this approach that need justification. The first is that cyber conflict is not narrowly defined by military use of cyber methods. A more complete understanding of cyber conflict involves the assessment of both cyber interests and cyber methods. The second aspect that needs justification is the use of the relationship between individuals and data as the defining factor of 'cyber'.

The justification of these points is arranged as follows. First, we ask the question of why one would not just accept the status quo and talk in terms of cyberwar. This discussion allows the failures of this narrow conception to be demonstrated. In the next subsection, *Terminology: Cyber*, the concept that cyber is best defined as the relationship between individuals and information is defended. The third subsection, *Terminology: War*, looks at the failings described impact on the current discussion and how a broader understanding of cyber conflict, and in particular talking in terms of conflict rather than war, leads to a more complete analysis.

Why not just cyberwar?

In general, I will use the term 'cyber conflict' in preference to 'cyber war'. In fact, both elements of the phrase 'cyber war' have their own challenges. The complications with 'war' are outlined below and the use of 'conflict' in preference is simply a means of sidestepping those complications and the baggage they bring with them. As will be shown, cyber conflict is a broader concept than cyber war and, as such, is more representative of the reality of modern conflict. 'Cyber' remains a poorly defined term and I defend a non-standard definition in which it is taken to reference the relationship between humans and information. Therefore, cyber conflict is a type of conflict that targets and/or uses the relationship between humans and information to achieve specific aims. As was stated at the start of this chapter, the type of conflict that is referenced is that which imposes harm. Large-scale conflict is the type of conflict that imposes harm on groups or populations rather than just on the individual.

The first step in defence of the previous paragraph, and particularly the modern definition of cyber conflict, is analysis of one of the first papers in the field. In the early nineties, John Arquilla and David Ronfeldt authored a paper (Arquilla & Ronfeldt 1993) entitled *Cyberwar is Coming!* It might be the first complete engagement with the idea of cyberwar. It is worth understanding what Arquilla and Ronfeldt were saying before seeing how the discussion progressed subsequently.

> The thesis of this think piece is that the information revolution will cause shifts both in how societies may come into conflict, and how their armed forces may wage war. (Arquilla & Ronfeldt 1993:27)

There are two key aspects to extract from that seemingly simple sentence. Firstly, and less contentiously, is the fact that we are undergoing an information revolution. This is also sometimes referred to as the digital revolution or, in the case of Lucian Floridi (2016), the 4th Revolution. This fact forms one of the 'axioms' of the project outlined in the prelude and no further discussion is possible here. It is enough to have the general understanding that the advent of computing, personal computing, big data, and the Internet have changed, and continue to change society in ways so significant that the result is best regarded as an upheaval.

The second aspect of the quote is that the changes of the information revolution will change how societies come into conflict and how they wage war. Given the first aspect extracted from the quote, this might seem uncontentious. As we will see, this

is not the case. There has been a significant amount of effort and thought dedicated to showing that either cyber methods are simply another method or weapon for the military to use, or that cyberwar is not even war, or both. A good starting point for the discussion of those issues is the sentence that follows the one quoted above,

> We offer a distinction between what we call "netwar"—societal-level ideational conflicts waged in part through internetted modes of communication—and "cyberwar" at the military level. (Arquilla & Ronfeldt 1997:27)

For readers thirty years later, 'cyberwar' has entered the popular lexicon, while 'netwar' has disappeared from our landscape. This has been a mistake. Of course, militaries all around the world have not been slow in understanding that cyber methods provide useful tools for achieving their ends. That form of cyberwar is clearly documented both academically and in more generalist books such as *Cyber War* by Richard Clarke (Clarke and Knake 2012) and *A Fierce Domain* by Justin Healey (Healey 2013). Documenting the history of cyber conflict is not the aim of this project and, in any case, is better left to those with 'inside knowledge' such as Healey and Clarke.

'Netwar' on the other hand is a term that has disappeared from our usage. It is puzzling as to why that is. The last thirty years have seen extensive development of "societal-level ideational conflicts waged in part through internetted modes of communication." The Internet has proven a savage battleground for informational conflict. Propaganda, electoral interference, governmental information campaigns, organised cyber trolling, battles over internet privacy, battles over net neutrality and many examples of cyber espionage are just some of the forms of ideational conflict that have appeared.

Why is concern with cyberwar dominant? Perhaps films such as *War Games* in 1983 helped popularise an idea of cyberwar as the primary form of informational conflict. Perhaps it is just more fun to think of the military use of cyber methods in which missiles are diverted and tanks disabled than to think of subtle and perhaps long-term ideational conflicts waged through internetted modes of communication. Military use probably does make a better plot for a novel or film.

The problem with the emphasis on cyberwar rather than netwar is that it neglects a significant section of how cyber methods can be used - it focuses only on a section of cyber conflict. On December 31st, 2016, Senator John McCain said, "When you attack a country, it is an act of war" (Schleifer & Walsh 2016). This seemingly common-sense view was stated in reference, not to a physical attack, but to the alleged activities of Russian hackers aiming to influence the results of the American elections. For the point that McCain makes to be coherent the concepts associated with netwar cannot be neglected.

Am I saying that we should resurrect the idea of netwar? It may be too late for that, but the use of the term 'cyber conflict' allows the inclusion of the type of conflict that is encompassed by netwar as well as the type encompassed by cyberwar. Cyber

conflict is a broader term that covers both types. As was stated in the opening paragraph of this chapter, the type of conflict that is referenced is that which causes harm. The focus of this project is on large-scale cyber conflict - that type of conflict that results in widespread harm to groups and populations.

There remains a certain amount of terminological difficulty in the following sections because other writers have used the term cyberwar with less precision. It would be complex to footnote every instance of this terminological difficulty. Where the difference is significant, the difference is highlighted. Where the difference does not have significance, it is passed over. The first step in clarifying the terminology is to have a clear understanding of what is meant by ‘cyber’.

Terminology: Cyber

I will use what is perhaps a non-standard definition of ‘cyber’ - that it denotes a relationship, or interface, between humans and information. This definition requires some discussion.

In 2016, in the first presidential debate (Golshan 2016), Donald Trump was widely regarded as having made a blunder when he talked about ‘the cyber’. It is pretty clear that most people do not talk in terms of ‘the cyber’ but it is less clear that there is either an accepted meaning or accepted usage for ‘cyber’.

We all have a pretty good idea about what it means when phrases such as cyber-crime, cyber-bullying or cybersex are used. It means that it has to do with the Internet, right? Despite the tempting simplicity of this definition of cyber, it is flawed for the reason that there are cyber actions that are carried out independently of the Internet. Take, for instance, a computer virus that is distributed on portable USB thumb drives. This virus is distributed independently of the Internet, but the process would still best be considered the use of cyber methods.

There is a vague, or less well-defined, definition of cyber that might be suggested - that cyber has to do with computers or, vaguer still, ‘computer stuff’. This is harder to dismiss directly due to its vagueness. However, as will be discussed more fully in subsequent chapters digital technology has become a major way in which we interface with information. As such, ‘computer stuff’ is simply the prevalent interface between humans and information. An interface is an implementation of a relationship. So, claiming that cyber refers to ‘computer stuff’ is a less precise way of making the claim that cyber refers to the relationship between humans and information within the current paradigm. It is this relationship that defines cyber rather than the specifics of computer hardware.

That cyber refers to a relationship is clearer if we talk, not in terms of computers, but in terms of information technology (IT) or information and communications technology (ICT). Information technology, IT, is a subset of ICT and talking in terms of ICT removes one of the problems that might arise with the term IT. As Wikipedia states (Wikipedia 2022b), the term IT is “commonly used as a synonym for computers and

computer networks." It might be easy for IT to be dismissed as a matter of hardware such as screens, printers, keyboards, computers, and mice. What is more important to this project is a conceptual understanding of the role the technology performs in allowing humans to interact with information. This interaction was described by Harold J. Leavitt and Thomas L. Whisler (1958) in one of the early analyses of IT as consisting of "three categories: techniques for processing, the application of statistical and mathematical methods to decision-making, and the simulation of higher-order thinking through computer programs". Whether or not one thinks that Leavitt and Whisler were correct in the details, it is clear that they did not believe that IT concerned only the hardware. Using the term ICT allows us to concentrate on the fact that the technology provides an interface between humans and information.

Leavitt and Whisler published their article in 1958. In many ways, it was highly speculative, and it was written prior to the huge impacts and social changes that have resulted from the digital revolution in general, and the development of the Internet in particular. One salient fact about ICT as an interface between information and humans that has become increasingly evident is that the interface is bi-directional. This means that the interface provides the means for humans to influence information, but also the means for information to influence humans. That the information with which we are surrounded affects our lives is hardly contentious in the 21st century. Information can have an instrumental influence on our lives but can also influence our beliefs and emotions - elements that we regard as key elements of our identity. These general observations were made concrete by a Facebook experiment reported in 2014 in which over half a million people were used as guinea pigs in an experiment to discover if Facebook could influence its users' emotions (Booth 2014).

ICT is an interface between information and humans. So, the claim that 'cyber' references the relationship between humans and information is equivalent to saying that it references ICT. Why not then say that cyber references information and communications technology rather than describing a relationship between humans and information? The first worry about using the term information and communications technology (ICT) is that it, and in particular the word 'technology', can still be interpreted as referencing only hardware. This particular interpretation does not suit the direction of this project.

The route that this project takes is to acknowledge the changes that the digital revolution is causing in society. These changes are influenced primarily by our relationship with information. The project is a conceptual analysis of that relationship and its effects on the theories we use to understand the ethics of conflict. Defining 'cyber' in terms that reference the relationship between humans and information allows an analysis of 'cyber' that is not reliant on the details of the technology itself. This project is not a technical analysis of the technologies of ICT. It is an analysis of the effects of the digital revolution on our increasingly important relationships with information and how those effects influence the ethical theories that we use in the context of cyber conflict.

Therefore, ‘cyber’ will be interpreted as referencing the relationship between humans and information. This is in no way incompatible with saying that ‘cyber’ concerns ICT in the broadest interpretation of ICT. What is emphasised here is the conceptual role that ‘cyber’ elements provide rather than the hardware or software of those elements. Critics may think that this definition is an abstraction of ‘cyber’, while advocates of the definition may feel that ICT is a particular instance of a relationship. This difference is moot in the context of this project. Talking in terms of a ‘cyber’ as a relationship allows the conceptual analysis that will be necessary without overestimating the importance of hardware.

Does that definition correspond with the real-world examples and usage with which we are accustomed in the context of cyber conflict? The definition demands that we interpret cyber conflict as conflict that is mediated through, or targets, the relationship between humans and information. Paradigmatic cases of cyberwar would, if that is true, be cases in which the relationship between humans and information technology is leveraged in order to create harm for political ends. In the two examples that follow, the harm and the political ends are, at this point in the discussion, taken for granted. The examples will demonstrate that it was the relationship between human and information that was targeted.

The two examples will be discussed much more fully in subsequent chapters but in the first, usually described under the term Stuxnet, America interfered with the control systems of the Iranian nuclear processing plant at Natanz (Vigliarolo 2017). The control systems provided information regarding the state of the processes involved - in particular, information regarding the state of a group of centrifuges used in processing uranium. In the system as designed, there was a relationship between the human operators of the processes and the information that described and controlled the processes. The attack interfered with this relationship. The primary aim of the attack was to interrupt the processes by corrupting the relationship of operators with the data they required. In this, it is a paradigmatic cyber-attack.

Sceptics will argue that the fact that it interfered with control systems is not what made Stuxnet a cyber-attack. What makes it a cyber-attack is that it was carried out using digital means - Stuxnet was, after all, a computer worm (Vigliarolo 2017). Stuxnet is doubly interesting because both the target and the method used in the attack, the attack vector, involved the relationship between humans and technology. The target did so because it was a control system. The vector did so because (again, as will be discussed more fully later) the distribution mechanism involved humans carrying the worm into protected areas on USB drives. Humans' relationship to information was corrupted to provide the vector. Stuxnet is a paradigmatic example of a cyber-attack in this respect.

The second example involves one of the early prosecutions of a cyber-criminal. Kevin Mitnick was arrested in 1995 and subsequently charged with multiple cases of wire fraud and computer fraud (Gengler 1999 & U.S. Government 1995). At one point (Markoff 1994), Mitnik was dubbed "cyberspace’s most wanted". However, he has always claimed (Simon et al. 2003) that his intrusions into computer systems were

based entirely on social engineering rather than computer skills. He managed to extract passwords and access from individuals and used those passwords to gain access. His methods were based on the interactions of humans and information rather than on specialist technical knowledge. His targets were largely informational assets of corporations - for instance, software repositories. By accessing that information, he corrupted the relationship that the owners had with that information. Unlike Stuxnet, it might be claimed that the vector - predominantly human engineering - is not part of an ICT in the narrow, hardware-based, definition of ICTs. However, it is based on the relationship between humans and information, and this fact supports the claim that 'cyber' references that relationship.

There is one more challenge to this interpretation of 'cyber' that is worth noting. If it is correct and 'cyber' references the relationship between humans and information, then any method that uses or interferes with this relationship can be designated as a cyber method. That raises the question as to whether such methods that occurred before the digital revolution should be designated as cyber methods. An example of such a method would be a propaganda campaign aimed at influencing the information available to a population. The challenge rests on the thought that this campaign, because it happened before the advent of ICTs, cannot be a cyber method. As will become clear, the route that this project follows is that the issues that surround cyber conflict are not novel but have been elevated from interesting bylines to core factors by the digital revolution. Yes, there were 'cyber' actions prior to the digital revolution because humans did have a relationship with information. The digital revolution has altered the value of our informational assets and the methods with which they can be targeted, and this has resulted in this form of conflict being significant in ways that it was not previously.

If this definition of cyber is coherent, how does this affect our understanding of the term 'cyberwar' and why would we prefer talking in terms of cyber conflict? The following section will outline the terminological problems of 'war' in the context of cyber conflict.

Terminology: War

In 2012, Thomas Rid published an influential paper titled, "Cyber War Will Not Take Place." (Rid 2012) It, as the title shows, was a response, at least in part, to Arquilla and Ronfeldt's earlier paper, and was rapidly expanded into a book of the same title (Rid 2013). Rid's work - particularly the book form - is a thoughtful and insightful analysis of 'cyberwar'. What has been most often noted in subsequent discussion and analysis is his claim that cyber methods would never meet the criteria that would allow them to be properly described as war.

> This article argued that the world never experienced an act of cyber war, which would have to be violent, instrumental, and – most importantly – politically attributed. No attack on record meets all of these criteria. Instead, the last decade saw increasingly sophisticated acts of network-enabled sabotage, espionage, and subversion. (Rid 2012:29)

In effect, it is taken that Rid used a Clausewitzian definition of war as being inherently violent, instrumental, and political (see Clausewitz 1976), to argue that ‘cyberwar’ was misnamed, because it had not, and would likely never, meet those criteria. In particular, it could never reach the level of violence that categorisation as war requires. It is a testament to how quickly the ‘cyber’ environment is changing that *Cyber War Will Not Take Place* is already, after less than a decade, a ‘historical’ work. With the benefit of hindsight, there are two aspects of the subsequent discussion that are worth highlighting with respect to Rid’s thesis.

The first is that from 2009 to 2012, a group of experts in international law had been invited to compile a study on how international law applies to cyber conflict. The resultant volume, titled Tallinn Manual on the International Law Applicable to Cyber Warfare (Schmitt 2013), was published in April of 2013. The word ‘violence’, which was critical in the previous paragraphs, is not a word that is generally used in legal works in this context. Rather, the Tallinn Manual considers whether cyber methods could ever be considered ‘a use of force’ which, for international lawyers, is a more tightly defined concept than violence.

The phrase ‘a use of force’ is a legal term of art with a great deal of importance in a number of fields. Its importance in international law and the law of armed conflict is encapsulated by the fact that the UN Charter prohibits, except in a very few circumstances, the use of force as an implement of international relations. UN Charter Article 2(4) states (United Nations 1945b),

> 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.

Exceptions are provided in UN Charter Article 51 (United Nations 1945a) which notes that nothing in the Charter “shall impair the inherent right of individual or collective self-defence if an armed attack occurs”. International law, in general, does not talk in terms of war and peace - even though the opening words of the UN Charter state that its purpose is to "maintain international peace and security". Rather, the prohibitions offered by international law are aimed at ‘the use of force’. In the legal context, a lot hinges on whether cyber methods may be regarded as the ‘use of force’.

The Tallinn Manual stated unequivocally that cyber methods might, in certain circumstances, be legally considered as the use of force. Section 11-6 of the Tallinn manual states,

> The International Group of Experts agreed, therefore, that any cyber operation which rises to the level of an ‘armed attack’ in terms of scale and effects pursuant to Rule 13, and which is conducted by or otherwise attributable to a state, qualifies as a ‘use of force’. (Schmitt 2013:47)

It would be handy for my argument to leave it at that. However, there was not unanimous consent to this in the Group of Experts, and it is noted that a likely

interpretation is that states will need to assess when a cyber action qualifies as a use of force. This approach will involve the analysis of the 'scope and effects' of an action.

> The approach focuses on both the level of harm inflicted and certain qualitative elements of a particular cyber operation. (Schmitt 2013:49)

The manual describes the possible qualitative concerns as severity, immediacy, directness, invasiveness, measurability, military character, and presumptive legality. This is not presented as a definitive list and many of the items have their own challenges.

The conclusion that emerges from the deliberations of these legal experts is that the use of cyber methods can amount to the use of force and at times can trigger the right of self-defence as described in UN Charter Article 51. They are quick to point out that not all use of cyber methods will meet these requirements (Schmitt 2013:49 §11.8 & §11.9). In a move that was prompted by an earlier paper by Michael Schmitt (1999), it was acknowledged that there is an "absence of a conclusive definitional threshold" when thinking about the application of 'use of force' to cyber methods, and a case-by-case assessment is likely to be required. That assessment will be of "both the level of harm inflicted and certain qualitative elements of a particular cyber operation" (Schmitt 2013:49 §11.9).

Rid's thesis was that cyberwar could not be considered 'war'. The opinions of the legal experts in the Tallinn manual stand at least slightly in opposition to that thesis.

The second aspect of the discussion that relates to Rid's thesis is that it is widely accepted that the nature of war is changing. For instance, Oxford University has a vibrant and active research group, the Changing Character of War Centre, which studies the issue. Philosophical papers have been written around the subject, notably "Should the Changing Character of War Affect Our Theories of War?" by Jovana Davidovic (2016). In sociological studies, Mary Kaldor (2017) tackles the theme in the influential book, *New and Old Wars.* Various treatments highlight different elements of this changing nature and attention has been directed at the increasingly asymmetrical nature of war, 'robot warfare' including the use of AI and drones, the increase in civil rather than international wars, the effects of technology and cyberwarfare, and various other themes. There is little doubt that warfare is changing more greatly now than it ever has before. Warfare has long been a technological undertaking and the pace of technological development is increasing.

Given that there is little doubt that the character of war is changing, is it appropriate to cling to our old understandings? Clausewitz wrote his unfinished masterpiece, *On War*, between 1816 and 1830. There is a strong line of argument that would suggest that any logic that was based on a definition of war that was 200 years old is inappropriate to the modern form of war. Possibly, the claim that war is necessarily violent, instrumental, and political needs to be viewed in this context. At this point, at the start of the project, the aim is only to emphasise that this understanding of war

can be limiting. Clausewitz's definition of war cannot be taken as an unchallengeable truth.

Rid's argument is not only that cyber actions do not amount to the 'use of force'. He also emphasises that cyber methods are often better categorised as sabotage, espionage, and subversion. This seems correct in most cases and is supported by the empirical work undertaken by Brandon Valeriano and his group (2018) which divides actual cyber actions into examples of disruption, espionage, and degradation. In the same way as the use of particular kinetic weapons might dictate the type of campaign, the use of cyber methods is most appropriate in certain types of campaign. Rid's categories of sabotage, espionage, and subversion certainly can be applied to cyber actions effectively and the categorisation is used as a tool in the following chapters. The division between these three categories is not always as clear cut as one might imagine because pigeon-holing cyber events in this way is often difficult, but the analysis in the following chapters will demonstrate a conceptual justification for these categories and the logical space that they define.

So, is 'cyberwar' war or not? That particular discussion can lead to a quagmire of semantic and theoretical bickering. In order to avoid that, one can take a hint from the Group of Experts. What they suggest is that in considering these issues in general one needs to think about "both the level of harm inflicted and certain qualitative elements of a particular cyber operation." This echoes the approach that is to be taken in this project. What is of concern, in moral terms, is not whether something is categorised as war, but the harm that the action inflicts. This approach is based on the premise that all large-scale conflict, i.e., large-scale harm creating actions and processes, has distinct ethical questions. In whatever way one divides the logical space of large-scale conflict, all the divisions will all share those ethical questions. However, it is the infliction of harm that gives an action moral importance and therefore it is the infliction of harm that is the focus of concern in this context.

It might be argued that this is an overly complex definition of the context. Surely, the argument goes, a definition such as that provided by the Correlates of War (Correlates of War 2022*)* and is discussed later would provide an easier understanding of the scope. The Correlates of War definition is based on a threshold of the number of deaths – war is a conflict that creates over 1000 deaths. Similar threshold approaches might be based on a particular level of harm, or a particular number of deaths, or another factor. The problem for this project with this type of threshold approach is twofold. Firstly, a hard-edged threshold does not accommodate the vaguer nature of warfare that is typical of modern conflict, in particular cyber conflict. Secondly, any threshold is purely arbitrary in nature. For instance, a campaign that created a great deal of non-lethal harm and 'only' 900 deaths would be excluded from consideration.

To a certain extent, threshold techniques are necessary for legal proceedings in which a binary decision is required, even though we may understand that the reality of the situation is messier and more complex. The aim of this project is not legal. Rather, it is to provide a framework in which the effects of cyber conflict can be understood in a moral context. That is inevitably going to be complex and while the utility of threshold

techniques is accepted in certain aspects of this analysis, any initial threshold that limits the scope of discussion is avoided. For the most part, from this point on, rather than using the words ‘war’ and ‘cyber war’, I will speak in terms of large-scale cyber conflict. This is a reminder that the presuppositions that words such as ‘war’ tend to import into the discussion should be avoided. It is not a comment that large-scale conflict is in any way a less serious matter than ‘war’.

Terminology: Final Thoughts

I have argued that large-scale cyber conflict is defined by two factors. Firstly, the fact that it is cyber conflict denotes that it relates to the relationship between humans and information. Secondly, it is large-scale conflict which means that it creates the imposition of harm on groups or populations. It is this widespread harm that demands an ethical analysis of the subject.

One of the ‘axioms’ outlined in the prelude was that **We are in the midst of a digital revolution.** The fact that we are in the midst of the radical changes that information technology is imposing on our lives and societies means that the final effects of the revolution are neither yet reached, nor is the final outcome of the revolution necessarily clear. The ultimate development of cyber conflict remains unclear and yet it is evident that cyber methods are capable of imposing significant harm. It is equally clear that those things that we value are being altered by the revolution. How those two factors, methods and interests, will ultimately interact remains to be seen. However, in the current environment certain characteristics of cyber conflict are becoming evident. In the following section, cyber conflict is shown to be persistent, dispersed, transversal and covert.

1.3.2. The Characteristics of Cyber Conflict

A fundamental fact regarding the type of conflict in which we are interested is that it involves methods for affecting other's interests negatively. The societal changes involved in the digital revolution have altered both the methods available and the interests that may be affected. It is the second aspect, that our interests are changing, that is sometimes ignored in the discussion of cyberwar. It is impossible to fully understand conflict without considering the contested interests as well as the methods by which they are contested. In the context of cyber conflict, these two elements will be termed cyber interests and cyber methods.

The transformations of the digital revolution are affecting the things that we value and the things that are significant in affecting our welfare. It will be shown throughout this project, but particularly in Chapter 3, that information or data that is associated with us is becoming increasingly significant in the provision of a meaningful and significant life. That body of information that is concerned with the individual forms a legitimate interest that has increasing significance in our lives. It is an informational interest or, probably more accurately, a group of informational interests because the information

is diverse and extensive. Such interests are termed cyber interests because they are informational assets with which we have a significant relationship.

The digital revolution is also creating new methods for affecting or influencing both cyber and other interests. Methods that leverage the relationship between humans and information are termed cyber methods. In most cases, as was discussed previously, these methods are conducted through, or by, ICTs. Often this means that the cyber method is conducted through the Internet, but this is not universal. There are some cyber methods that are independent of the Internet.

The history of recorded cyber methods is relatively short, and the characteristics of cyber conflict are emergent rather than firmly established. Hence, the four characteristics described are neither intended to be an exhaustive list nor a final one. Rather, they are an overview of the characteristics of cyber methods - and the terminology associated with these characteristics - that are relevant to the discussion in the following chapters. The four characteristics that are outlined below are transversality, covert nature, dispersed nature, and persistence.

Transversality

Transversality is a way to describe the fact that cyber methods can have results and consequences in both the digital and the physical domains. If I were a nefarious hacker, I might access your computer in order to destroy data - which would be an example of a cyber method having an effect in the digital domain. Alternatively, I might use my unauthorised access to your computer to overload the power supply to your computer and create a fire - which is an example of a cyber method having a physical effect.

Ransomware is a typical example of cyber methods targeting cyber assets and is a growing industry. Estimates in 2016 put the number of daily ransomware cases at about 4000 (U.S. Government 2016a). It is likely that that daily number has increased dramatically in the intervening years. Its mode of action is to deny victims access to their own data unless a fee is paid. The WannaCry attack in 2017 shut down hundreds of thousands of computers around the world. Users were locked out from their own machines and presented with on-screen demands for payment to remove the block. The UK health service was particularly hard hit with an estimated 1 percent of all care impacted for a week at the cost of £20 million, a price that does not include the forensic and update costs subsequent to the attack.

The logic of ransomware attacks is straightforward. The cyber interests of the users are of value to the users and in many cases, users will choose to pay the ransom rather than lose the data. In financial terms, ransomware demonstrates the value of information. Payment also happens in cyberspace these days. Whereas a traditional ransom might be paid in unmarked banknotes, ransomware ransoms are typically paid in Bitcoin or other cyber currencies.

While ransomware is an example of informational assets being targeted by informational means, cyber methods can also be used to target physical assets. The most widely used example of cyber methods having effects in the physical domain is Stuxnet. It was a cyber action by the U.S. government that successfully targeted the Iranian nuclear programme. The result of the cyber intrusion was the malfunctioning and subsequent destruction of a significant percentage of centrifuges used to process uranium. Estimates vary but the usual guess is that Stuxnet set back the programme by a number of years. It makes a fascinating case study and will be revisited in depth later.

Using transversality to describe this feature of cyber methods follows the usage of Mariarosaria Taddeo. She uses the term informational war, IW, to describe conflict involving cyber methods,

> Nevertheless, the deployment of ICTs gives rise to a completely new form of warfare, whose main peculiarity is that of being transversal. IW is transversal with respect to the environment in which it can be waged, the kinds of agents involved in it, and the modes of combat. Such transversality represents the ultimate difference between IW and classic war, and it is the aspect of IW from which policy-related and ethical problems arise. (Taddeo 2012:112)

Taddeo uses the term transversality in a slightly broader way from that in which it is used in this project. According to Taddeo IW is environmentally transversal due to the fact that it can originate and be effective in either (or both of) the physical or digital domain. It is transversal in the kinds of agents involved because it does not necessitate military involvement. It is transversal in its ‘modes’ for two reasons according to Taddeo. The first is that it concerns a wide range of techniques ranging from Denial of Service (DDoS) type attacks to the use of computerised autonomous weapons. The second is that it ranges from the ‘soft’ non-violent actions such as voter manipulation to extremely violent actions such as the destruction of a power station. Although Taddeo’s analysis does not conflict with the points made here or subsequent analysis, I prefer to use transversality to refer only to the boundary crossing between the physical and digital. The range of effects of cyber methods is a separate issue.

Transversality raises an interesting point. Can a physical action have informational effects? The answer is that it can. For instance, bombing a datacentre is a physical action that evidently has physical consequences. It also has informational consequences that may echo through the digital domain. Data and processing may be destroyed and that is an effect in the digital domain. Transversality is not unique to cyber actions, but it is a characteristic of a large number of cyber actions.

Covert

It is sometimes noted that cyber actions are often covert. Looking up the word ‘covert’ in the Oxford English Dictionary, one finds that the most appropriate definition is "Concealed, hidden, secret; disguised." There is some discussion in cyberwar literature about what covert means with regard to cyber methods. Justin Healey

(2012:75), in A Fierce Domain, comments that “The most dangerous trend today is the apparent willingness of nations to engage in covert campaigns against each other.” Attribution of responsibility for cyber-attacks has been seen as a problematic issue. It is intuitive to think that digital actions do not carry the same attributable stamp as physical actions. If I write you an offensive letter, you can check the sending postcode, my writing and other attributes of the letter and be sure that the letter comes from me. Digital communication is normally attributable in the same way by experts. For the inexpert, attribution is challenging. Even with expertise, it is easy to imagine that methods exist by which attribution might be avoided.

There is an image of an anonymous hacker - probably in a darkened room surrounded by pizza boxes - who by using digital methods that are opaque to normal people finds out about our lives and interferes with them. This image is perpetuated in films such as War Games, The Net, and many others. The anonymous hacker is an established meme. In fact, one of the most notorious hacking groups is called ‘Anonymous’ (who rather ironically campaign, among other things, for accountability in politics). This icon of the hacker is not without substance. In 1998, members of Congress met with members of the L0pht, a hacker collective based in Boston. Although this has been seen as a crucial step in moving governmental understanding forward, the appearance of the seven members of L0pht - who at the hearing used the names Mudge, Weld, Brian Oblivion, Kingpin, Space Rogue, Tan, and Stephen Von Neumann - did little to dispel the image of the archetypal hacker.

In fact, there is, or possibly only was, a romanticised vision of this hacker as a modern day outlaw. In 1998, following the SOLAR SUNRISE attacks - originally feared to be related to the ongoing conflict with Iraq and later attributed to two Californian teenagers with the help of an Israeli ‘mentor’ - Israeli press reported that “People see him as the outlaw of our time, and they really like the fact that this little Israeli went up against the big guys - the Pentagon” (Healey 2012:133). Likewise, when Kevin Mitnick, who we met earlier, was arrested in 1995 there was a popular campaign to ‘Free Kevin’. According to Chris Painter (Healey 2012:83), who was in charge of the case against Mitnick, a plane circled the courtroom with the ‘Free Kevin’ message on a banner. Whether that romanticised vision of the hacker as a type of Robin Hood remains today is debatable but it certainly has been a thread in the popular perception of cyber conflict.

The offensive cyber teams of nation states are largely classified, but they definitely exist. The most famous cyber offensive, Stuxnet, was intended as an unattributable covert operation and it was only through accident and coding errors that the U.S. eventually had to accept responsibility. Few doubt that many of the claims of Chinese digital espionage that the Trump government rails over are justified. It is clear that North Korea, Iran, Russia, Great Britain, and others all have offensive cyber programmes. A great deal of this activity is classified. There is no reason to believe that nations are not investing heavily in offensive cyber capabilities and there is no reason to believe that successes are not being achieved.

So, what is the disagreement relating to the use of ‘covert’ in relation to cyber methods? The issue really comes down to what is implied by the term. There is a strong interpretation of ‘covert’ that implies that the action is actively hidden, and the actor aims to remain unknown or anonymous. There is a more modest interpretation in which the action or actor is simply not overt.

There are certainly cases that are not covert by either interpretation. Valeriano and his colleagues (Valeriano et al. 2018), in their analysis of cyber strategy have a focus on interstate coercion and whether cyber methods are effective methods of coercion. They note that for coercion to be effective, it is often counterproductive if the opponent is truly unaware of the source of the attack and signalling often forms a core element of aggressive actions. Jonathan Diamond notes (Healey et al. 2012:150) that what is described as patriotic hacking - non-military hackers acting on behalf of their nation - demonstrates that “though often cited as one of its greatest difficulties, the attribution problem is not the *sine qua non* of cyber conflict.” There is a range of cyber methods that are overt.

A recent study based on state-sponsored cyber actions in 2016, 2017 and 2018 found that,

> First, the vast majority of incidents (70, or 85%) resulted in some form of public attribution, with only 12 incidents (15%) not being attributed to a perpetrator. A small number of incidents, 7 (9%), included attributions involving both government(s) and private actor(s). (Kuerbis 2018)

These figures might suggest that the idea that cyber actions are covert is not significant. However, certainty in the field of covert operations is difficult. It is entirely possible that there were numerous successful incidents that were not only not attributed but not even noticed. The most successful covert cyber actions will by definition not appear in this type of figure.

The strong interpretation of covert is largely untenable. Most cyber security experts agree that attribution is possible in most cases. As Valeriano et al. claim,

> In fact, the idea of secret cyber operations is often overstated. A cyber action, by definition, can be witnessed and observed by any capable adversary. These actions are covert in that they are deniable (Carson and Yarhi-Milo 2017:128). States leverage cyber tools because they can avoid responsibility for their actions and generally avoid attributing their operations for days, if not years, after the event. (Valeriano et al. 2018:15)

The forensic effort and costs may not be worth the return but in most cases, it is possible to know from where the attack originated with a reasonable degree of certainty. On first inspection, and to less technically adept individuals, it may appear that the originators are hidden, but the evidence does in fact exist. What the initial uncertainty provides is not so much secrecy as plausible deniability.

The more modest interpretation of ‘covert,’ in which it is used to imply the opposite of ‘overt,’ is more appropriate in the context of cyber actions. Cyber methods are not by their nature overt in the way that most conventional weapons are overt. It is possible to make the actions overt by advertising them as is done in the case of ‘patriotic hacking’, but their nature is that the methods do not advertise themselves. This fact can be leveraged to provide either strongly covert actions, deniable actions, or both. That is the form of ‘covert’ that is applicable to cyber actions. They are not intrinsically overt is a more accurate way of describing this characteristic. ‘Covert’ however is easier on the tongue.

Dispersed

The idea of ‘patriotic hacking’ is symptomatic of a further characteristic of cyber conflict, its dispersed nature. Both the agents responsible for cyber acts and the harm inflicted on its targets can be more dispersed than in traditional conflicts. The war in Kosovo provides an example of patriotic hacking. Kenneth Geers, U.S. Representative Cooperative Cyber Defence Centre of Excellence, describes the events,

> As NATO planes began to bomb Serbia, numerous pro-Serbian (or anti-Western) hacker groups, such as the “Black Hand”, began to attack NATO Internet infrastructure. It is unknown whether any of the hackers worked directly for the Yugoslav military; regardless, their stated goal was to disrupt NATO’s military operations.
>
> The Black Hand, which borrowed its name from the Pan-Slavic secret society that helped to start World War I, claimed it could enumerate NATO’s “most important” computers, and that through hacking it would attempt to “delete all the data” on them. The group claimed success on at least one U.S. Navy computer which it claimed was subsequently taken off-line. (Geers 2009)

The hackers here were not part of the state military. This was noted as a developing trend in cyber conflict and prompted the development of a spectrum of state responsibility for cyber-attacks which provides a useful tool for understanding how actions may be related to the state (Healey 2012:2),

> The Spectrum of State Responsibility
>
> 1. State-prohibited. The national government will help stop the third-party attack.
>
> 2. State-prohibited-but-inadequate. The national government is cooperative but unable to stop the third-party attack.
>
> 3. State-ignored. The national government knows about the third-party attacks but is unwilling to take any official action.

4. State-encouraged. Third parties control and conduct the attack, but the national government encourages them as a matter of policy.

5. State-shaped. Third parties control and conduct the attack, but the state provides some support.

6. State-coordinated. The national government coordinates third-party attackers such as by “suggesting” operational details.

7. State-ordered. The national government directs third-party proxies to conduct the attack on its behalf.

8. State-rogue-conducted. Out-of-control elements of cyber forces of the national government conduct the attack.

9. State-executed. The national government conducts the attack using cyber forces under their direct control.

10. State-integrated. The national government attacks using integrated third-party proxies and government cyber forces.

Jason Healey - who went on to author/edit *A Fierce Domain* - created this list when he was with the Joint Task Force for Computer Network Defence between 1998 and 2001. It demonstrates the divergence of responsibility from the conventional model of harm-inducing actions being the provenance of the military. The relationship between governments, militaries and individuals is complex and diverse and the actors in offensive cyber conflict are not limited to a single organisation such as 'the military'. In this they are more dispersed than the actors involved in conventional war.

There are strong reasons to believe that increasingly the targets of cyber conflict will be equally dispersed. Conventionally, at least in theory, the target of military action has been predominantly other militaries. There are, of course, exceptions to this such as the allied bombing during WWII but the paradigm of ‘war’ is that militaries fight other militaries.

The practicalities of cyber conflict mean that a focus on actions against military forces is not necessarily the most efficient course and this lack of efficiency may result in a tendency for more dispersed actions. A substantive difference between cyber and conventional methods is that most forms of cyber-attack rely on vulnerabilities in the target systems.[15] Military systems are isolated and hardened against intrusion. That is not to say that they cannot be compromised. Such compromises have been demonstrated in the real world (e.g., Weinberger 2007). However, a cyber-attack of

[15] The exceptions are denial of service type attacks which just swamp systems and societal attacks that rely on information manipulation on open platforms.

that sort is likely to be a ‘single-use’ deployment. It will rely on a vulnerability that is unknown to the defenders. Its use by the attackers is likely to reveal it to the defenders who will be able to protect against further compromises by this vector. It is probable that military systems will become increasingly hard to compromise. In comparison to other systems, they have two advantages. They can be isolated from the Internet and other communication, and they have a degree of top-down control which allows security measures to be enforced.

In contrast, the general, publicly accessible Internet does not have those advantages. There is extremely little top-down control, the general implementation of security measures relies on the actions of individuals, and connectivity is an inherent characteristic. These factors are aspects of the Internet that are regarded by many as essential to personal liberties, or Internet 'freedoms'.

There is a wonderful picture of the L0pht hacking group returning to congress in 2018, twenty years after their first appearance which put the cat among the pigeons regarding cybersecurity.[16] It is wonderful partly because the ‘outlaws’ look nothing like the young men who scared congress all those years ago. Rather, they look like established system security experts - which is exactly what they now are. Among the warnings that they gave on both occasions are the fact that there is little top-down control over current public systems and that those systems remain constructed on infrastructure that was not designed for security.

What that means is that vulnerabilities on the Internet are much more plausible than on military networks. This results in the fact that an easy way of applying political pressure to a nation is to apply very widespread pressures to that nation’s citizens. This is the form, sometimes known as hacktivism, that is likely to become increasingly prevalent in cyber conflict (see Lucas 2017). Highly discriminate attacks are likely to become increasingly difficult, in which case highly indiscriminate attacks are increasingly likely.

We may not have reached the point when the security around critical infrastructure is acceptable. It may still be possible to compromise military operations by cyber means. However, there is a pattern of increasingly successful defensive cybersecurity around military and large civil infrastructure. It is likely that the targets of cyber conflict will increasingly be dispersed - they will be the populations rather than the militaries of nations.

A developing characteristic of cyber conflict is that it is dispersed, both in terms of the actors and the targets.

[16] See, for example, Fisher 2018, for a history of the L0pht group.

Persistent

The final characteristic of cyber conflict, persistence, is closely related to this dispersed nature. Persistence implies that cyber conflict is characteristically ongoing and best regarded as a consistent rivalry. Its intensity may fluctuate but it is always there.

Of course, cyber methods are capable of spectacular one-off events. Brandon Valeriano has colourfully called these events the 'megafauna' events and the commonly used example is Stuxnet - which will be discussed in more detail later - in which the U.S. launched an extended attack on Iran's nuclear programme. This type of event has been scrutinised in some detail due to the fact that if any cyber actions will meet the legal definition of 'armed conflict', it will be these. This disproportionate scrutiny is misleading.

One of the first attempts at an empirical analysis of cyber conflict is *Cyber Strategy: The Evolving Character of Power and Coercion* by Valeriano et al. On reading it one emerges with an understanding that the authors believe that cyber methods are not best categorised as these one-off events,

> We find that the utility of cyber strategy is as a form of political warfare optimized for the 21st century that relies on tacit bargaining and ambiguous signaling to help rival states achieve a position of relative advantage in long-term competition. (Valeriano et al. 2018:13)

The authors are clear that there are challenges in empirical studies of cyber methods,

> While most military plans involving cyber attributes remain highly classified, the authors piece together strategies based on observations of attacks over time and through the policy discussion in unclassified space. (Valeriano et al. 2018:Abstract)

The lack of easily accessible historical data is one challenge. Another, that is not overtly noted, is that the nature of cyber conflict is changing rapidly. Rather than having had a digital revolution, we are in the midst of it. In many ways, we are still emerging from a period where the potential of digital communication overshadowed the necessity for cyber security. Our use of the digital realm and our idea of what digital safety entails remain immature, but they are rapidly maturing.

However, there is a pattern that is becoming evident. Megafauna events are expensive, difficult and, most likely, in contravention of international law. Moreover, as Valeriano argues (Valeriano et al. 2018:13), they are largely ineffective as stand-alone instruments of coercion. These authors operate in a framework that focuses on international coercion and coercive diplomacy. A potential criticism of their impressive work is that they neglect the longer timescale. Although, the premise of their analysis of cyber conflict is that they can assess the efficacy of cyber conflict by its success at

coercion in gaining "relative advantage in long-term competition." It is possible that we do not yet have the perspective to do that in the long-term.

Either way, the pattern is that cyber conflict forms a part of international relations that is ongoing rather than exceptional. The public may only be aware of the large events but campaigns of sabotage, subversion, and espionage - to use Thomas Rid's designations - are going on all the time.

Richard Harknett, an academic who has had an evident impact on U.S. policy, promotes the understanding of persistence in cyber conflict and has argued that U.S. policy must incorporate that understanding,

> U.S. national security, advancement of interests, and the development of international norms require persistent cyber engagement, not operational restraint, in an environment of constant activity. (Fischerkeller & Harknett 2017:382)

The U.S. National Cyber Strategy document claims that "This now-persistent engagement in cyberspace is already altering the strategic balance of power." (U.S. Government 2018)

It might be argued that 'persistence' is a current catchphrase in the discussion of cyber security. However, it does describe concisely that the scope of cyber conflict is such that, while cyber methods are capable of singular events, even those singular events are best conceptualised as exceptions to a consistent and constant struggle between rivals. Sometimes that struggle may be best viewed as a jostling for advantage, sometimes it may be best viewed as overt and attributable conflict, and only occasionally are there instances of the struggle that are these singular events. What all these scopes have in common is that they inflict harm. This harm may be used either for immediate effect as a form of coercion, or for longer-term effect as part of an extended struggle for advantage.

Cyber Interests and Methods

In conclusion, cyber methods are a recent phenomenon, and their use is developing. However, there are patterns that are observable, and these patterns demonstrate the key characteristics of cyber conflict. Those characteristics are its transversality, its covert nature, its persistence, and its dispersed nature. They form a general understanding of how best to conceptualise cyber conflict rather than a set of hard and fast rules, but this is appropriate given the rapid development and fluidity in the field.

1.4. And so …

The theme of this project is the effect of cyber conflict on Just War Theory. The aim of this section was to map out the landscape of the discussion.

(1) Just War Theory is a moderate theory of war.
(2) Moderate theories of war allow the infliction of harm in specific circumstances.
(3) Implementation of a principle of proportionality, a principle of discrimination and a principle of just cause are necessary for a successful ethical theory of large-scale conflict.
(4) The term ‘cyber’ refers to the relationship between humans and information.
(5) Cyber conflict is a type of conflict that targets and/or uses the relationship between humans and information to achieve specific aims.
(6) We are involved in a digital revolution that is changing the character and nature of our society and will inevitably influence the manner and means with which we can inflict harm on others. The new valued elements of society are termed cyber interests. The new means of inflicting harm are termed cyber methods.
(7) Cyber interests are defined as informational interests that are relevant to human lives. Those interests include families of cyber interests such as personal data control and public data access.
(8) Cyber methods can be used both in military contexts (cyberwar) and in non-military contexts (netwar). Both contexts need to be acknowledged and included in any complete analysis. Cyber methods are characterised as being transversal, covert, dispersed, and persistent. These provide general descriptive characteristics rather than a tightly formed definition.

In general, I am going to avoid the word ‘cyberwar’ and use the term large-scale cyber conflict. This is to avoid the presuppositions that might be imported with the term ‘war’.

Because JWT aims to compare harm and benefit, a form of quantifying both those harms and those benefits is necessary. The next two chapters look at how we might conceptualise that quantification, both in general and with regard to large-scale cyber conflict. After that, we will be in a position to analyse the effects of cyber methods and interests on moderate theories of war.

2. Against Body Count Morality

Cyber attacks are a threat not only to sophisticated information technological systems, but also to a community as a whole.... The threats posed by cyber warfare have often been underestimated since, fortunately, they have so far not resulted in the loss of any lives.

Toomas Hendrik Ilves (2007)

2.1 Introduction

It has been argued that a significant reason that an ethical analysis is appropriate to the study of large-scale conflict is the widespread infliction of harm. The purpose of this chapter is to analyse the main method of assessment of harm in current theories of large-scale conflict. In these theories, harm is commonly represented by the number of deaths. It will be shown that this representation is not adequate for a framework that serves our needs in assessing cyber conflict.

The core thesis of this chapter is that other factors matter as much, and in certain circumstances more, than death. On a personal or individual level, this is hard to contemplate because of the final nature of death for the individual, but on a societal level, it is a largely ignored truth. When this chapter was first written the idea of a global pandemic was the domain of sensational movies and television series. A year later, the zeitgeist is different. The pandemic, and our emergence from it, have forced people to come to grips with some troubling ethical problems such as ‘how much economic harm might justify the loss of a hundred lives?’

They are troubling because the answers are often uncomfortable. They go against the grain. For at least some people, there is a fundamental belief, or desire, that human life should be sacrosanct; that the loss of even a single human life should be an unjustifiable evil. This is the grain that the moral quandaries of the pandemic cut across. It appears that when push comes to shove, we do not adhere to that belief very strongly. It has not taken much for there to be a widespread acceptance of the fact that a considerable number of deaths may be justified by economic recovery. It is at least plausible that death, or a number of deaths, can be justified by other factors.

Of course, this was evident before the pandemic. Speed limits on highways and motorways are a case in point. The evidence shows that the higher the limit the more deaths will result. If we elevate deaths to a position of importance by which it becomes the only important criteria of moral evaluation, we would have extremely low speed limits, or not be able to drive at all. And yet this is not the decision that has been made by governments around the world. The functional benefits of motor traffic have been

balanced against the lives that it costs, and appropriate speed limits have been constructed.

The area of investigation of this chapter is whether death is an adequate metric for evaluating harm in large-scale conflict. Without wishing to spoil the surprise, the conclusion will be that it is not. Using death as the metric, an approach that I will term body count morality (BCM), not only does not reflect the true harms involved in large-scale conflict but also does not reflect the assessments we routinely make in other areas. Body count morality distorts moral evaluations, and this is seen particularly clearly in cases where the harms created are various and across multiple aspects of our lives. Moreover, the defence that body count is simply used as a rule of thumb, or a heuristic device, does not stand up well to investigation.

This truth, that body count morality distorts moral evaluations, is particularly relevant to discussion of cyber conflict. The persistent and dispersed nature of cyber conflict described in the previous chapter results in the fact that many of the effects of cyber methods are not necessarily loss of life. As will be argued in this section, that does not make them any less significant. In particular, the concept that cyberwarfare is in some way perfect, ideal, or harm-free is simply wrong.

Whereas the discussion of terminology that was carried out in Chapter 1 was largely conceptual, here it is practical and pragmatic. The most widely used definition of war is provided by the Correlates of War project and is discussed more fully in Section 2.4. That definition includes the stipulation that war accounts for over one thousand deaths. This arbitrary stipulation defines the landscape of the discussion. Talking about these moral issues objectively and impartially while using terminology that assumes that number of deaths is an adequate metric is impossible.

The use of body count morality will be shown to be widely prevalent in the study of large-scale conflict. There is a certain irony in this. We know that some individuals go willingly into conflicts that may result in their own deaths. Some combatants may be coerced, some may be in denial, but there are those who believe that the benefits of the action are worth sacrificing their lives. In entering into this type of lethal conflict, individuals and groups are inherently accepting that there are interests that are more valuable than lives. One of the fundamental motivational tenets of ‘war’ is that there are interests that are worth dying for.

2.2 Motivation

One does not need to be a pacifist to believe that in large-scale conflict the imposition of significant harms including, but not limited to, death is at the very least undesirable, and probably impermissible in many circumstances. An argument in favour of the use of force in these situations is that the harm caused may be justified by their contribution to morally preferable outcomes. However that argument is framed - as lesser evil, greater good, through the doctrine of double effect - there nonetheless is

an understanding that there are strong moral limitations on the nature and extent of harm to be imposed, regardless of the potential justification.

A moderate theory of conflict must include principles that aim to limit the excessive imposition of harm. For example, in JWT, these principles include proportionality and necessity. Both principles place limits on the amount or extent of harm that can be imposed on one side by the other. The limit created by the principle of necessity is grounded in the minimum harm that is required to achieve the aim. Options that involve more harm than this minimal extent of harm are impermissible. Proportionality balances the harm that is to be avoided against the harm that is to be used for that avoidance. So, if the harm to be avoided is extremely minor, it is demanded that one avoids actions that cause major harm.

Inherent in both these principles is an understanding that there is a method by which the harm of actions can be quantified. This quantification of harm is intrinsic to any theory that holds that evaluating the permissibility of conflict involves weighing the harms caused against the harms prevented. The process of quantification of harm is problematic in all ethical studies. How can one quantify a multivariate and complex phenomenon such as harm in the context of large-scale conflict?

Because harm is a complex issue, one way to approach this problem is by the abstraction of the harms in order to make the quantification possible and one prevalent way of performing that abstraction is to reduce consideration of all harms to consideration only of killing. I term this body count morality (BCM).

It is worth mentioning early on that BCM is only plausible in the context of large populations. On the individual scale, it is implausible. Abstracting all harms to life/death on the personal level would result in a framework in which as long as you were alive then no harm had been done. This is clearly untrue. One has other significant and important aspects of life that need inclusion in any assessment of overall harm, even of significant harm. This glimpse of the impracticality of individualistic body count morality hints at the overall problems to come.

It may seem unlikely that anyone might subscribe to this form of moral calculation. In Section 2.3 I outline how widespread and pervasive BCM is. It is a reduction of the complex significance of the conflict to a simple metric - the number of deaths that the conflict directly involves. A simple count of the bodies left on the battlefield is used as the single, or at least dominant, metric to assess the effects and impact of actions. In attempting to provide a quantifiable metric, this reduction eliminates the consideration of significant other forms of harm and contributes to the creation of problematic moral blind spots in which significant harm exists that is not featured in moral assessments.

Section 2.4 provides an analysis of the manner in which BCM is used as an abstraction to define conflicts and simplify the arithmetic of ethical calculation. Body count morality treats the quanta of ethical calculation as a single life and therefore the

calculation deals only with whole numbers and those whole numbers can be easily tallied in ethical calculation.

Of course, this may just be a technique or generalisation that is used to quantify a phenomenon that would otherwise be difficult to quantify. Perhaps body count is simply a rule of thumb, or a heuristic device, by which the complications of assessing harm can be mitigated? In Section 2.5, I argue that this defence of BCM cannot stand.

This critique of BCM is only significant if there are significant harms that are left unconsidered. If the use of BCM is inappropriate, it must be demonstrable that significant harms are systemically overlooked by its use. Left to last in Section 2.6 but possibly of greatest importance is the description of what types of harm are neglected by BCM and how this neglect is effected.

Overall, the implicit claim that is made in body count morality is that killing is all that matters. In rejecting this claim, one might adopt a range of other positions. It is helpful to itemise some ways in which the claim might be presented. In terms of the assessment of the harm of conflict,

A. Killing is the only important moral factor
B. Killing is the most important moral factor
C. Killing is one of the important moral factors
D. Killing is not one of the most important moral factors
E. Killing is not a moral factor

The target of the chapter in broad terms is to evaluate which of these claims is tenable. It will be assumed that option C is uncontroversial while both options D and E seem to devalue killing excessively. More specifically then, the target is to show that claim A cannot be tenable.

2.3 The Prevalence of Body Count Morality

It might feel as though this talk of the dangers of body count morality is sensationalist. Surely, the truth is not that bad. Sadly, one does not have to venture too far into philosophical literature to be disavowed of that hope.

Jeff McMahan, whose seminal book is called *Killing in War* (McMahan 2009), is one of the dominant voices in the philosophical study of war who has famously argued that the traditionalist interpretation of JWT is misguided. His fascinating 2015 paper, Proportionality and Time, is a good example of the tendency toward body count methodology. Although early in the paper (McMahan 2015:6) the argument is framed in general terms of “proportionality in harms caused to people who are not liable to

those harms," when the argument becomes more detailed (McMahan 2015:6) a body count form becomes apparent.

> Suppose that the correct judgment in a particular case is that the achievement of the just cause can justify the killing of 1,000 innocent people (by which I here mean "people not liable to be killed") as a side effect of military action. At this point, before the initiation of the war, all the available evidence indicates that the just cause can be achieved without killing more than 1,000 innocent people.

From that point onwards in the paper, the discussion is framed in terms of deaths. Body count methodology has triumphed. I am not commenting on the overall content of the excellent paper but highlighting the drift into body count morality.

McMahan is far from alone in this tendency. At least part of recent thinking surrounding war by moral philosophers has been based on philosophical investigation of self-defence on the personal level. These investigations have, almost without exception, focused on the defence of one's life. Domestic violence, rape, beatings, fistfights, bar brawls and other forms of non-lethal harm are all rare in these discussions whereas killing is common.

It is possible that this tendency stems from a corresponding tendency in the discussion of personal self-defence. Judith Thomson (1976) phrases one of the questions of personal self-defence in these terms, "surely it is permissible to kill a man if that is the only way in which you can prevent him from killing you!". Michael Otsuka (1994) starts his paper with the description, "Many philosophers subscribe to the common belief that you are morally permitted to kill a person who endangers your life whenever such killing is necessary to prevent yourself from being killed." In the legal literature surrounding self-defence, George Fletcher (1973:371) wrote what became an influential paper on self-defence that used the memorable example of the psychotic attacker, "Imagine that your companion in an elevator goes berserk and attacks you with a knife. There is no escape: the only way to avoid serious bodily harm or even death is to kill him." In this example, the options available to one are reduced to lethal force to abstract away some other complications. In general, philosophising about self-defence has involved discussions of lethal force and balancing the resultant body counts. A quotation from Thomson (1991:298) demonstrates both the tendency towards body count methodology and the type of calculations the method allows,

> In the cases we looked at in Sections I through IV, what is in question is one life for one: yours and that of a person Y whom you have to kill if you are to save your life. We will look briefly in Section X at cases in which several lives would be saved by the killing of Y; let us consider here only cases in which what is in question is one life for one.

This tendency also extends to those theorists more directly concerned with war and conflict. The thought is clearly stated by Leonard Kahn (2018:13) when he says, “It is killing and death that make war unique—and uniquely terrible.” The acceptance that violence and killing are characteristic and uniquely morally relevant elements of war is widespread in the literature. Often descriptions of war are phrased in ways that assume that this is the case. Michael Walzer (2006:41) says that there are two clusters of prohibitions in the morality of war, "The first cluster specifies when and how they can kill, the second who they can kill." David Rodin (Rodin 2004:63) describes the challenge of self-defence in war as "why it is that defenders are morally entitled to kill aggressors in situations of self-defense." Helen Frowe (2015:173) describes the reductivist view of war as holding, "that the rules governing killing in war are simply the rules governing killing between individuals." Jeff McMahan (2009), as mentioned earlier, titles his book, "Killing in War". Ryan Jenkins and Bradley Strawser (Jenkins et al. 2017) title theirs, "Who should die? The ethics of killing in war." Further examples are easy to find and the fact that the language that is used when discussing the morality of war centres around killing demonstrates that war is characterised predominantly in terms of killing.

A prominent example in the discussion of the ethics of war is the ‘argument from political aggression,’ also referred to as the ‘bloodless invasion’ example. The underlying question that its variants pose is whether a defender is permitted to respond with lethal force to an attack that does not threaten the use of force. The form of the argument was made mainstream by David Rodin (2002) who used to challenge the right to national defence. If the right to national defence is grounded in rights to security, Rodin argued, then cases in which lives were not threatened would not trigger a right to self-defence, but we traditionally believe that defensive force can be used to prevent annexation. In the discussion around the argument from political aggression, it is often claimed that war is impossible without killing - that the defence will inevitably cause the loss of life. In his treatment of the argument from political aggression, Seth Lazar (2014) acknowledges this claim and notes that he believes it is a practical truism that even a defensive war cannot be undertaken without killing. He goes on to say that denial of this fact would diminish the understanding of the moral seriousness of war.

However, this demonstrable tendency to talk in terms of body count does not show that philosophers are unaware of the other forms of harm that are involved in war. In fact, particularly in the discussion of bloodless invasions, it is a common technique to divide an individual's interests into vital and non-vital categories. As David Rodin (2014:80) says,

> By ‘vital interests’ I mean those centrally important interests, the unjust threat to which can justify lethal force in a domestic context of self-defence. These are in broad terms: threat to life, substantial threat to bodily integrity (including

> loss of limb, torture, and rape), profound attacks on liberty such as slavery, and permanent or long-standing displacement from one's home.

Rodin is not alone in talking of vital interests. Saba Bazargan-Forward (2017:142) says,

> Crucial to the concept of a bloodless invasion is that the victims' vital interests – viz., their lives and their bodily integrity – are conditionally threatened as a means to undermining their non-vital interests – viz., their political rights.

There are two ways in which the idea of vital interests might be conceptualised; either as a threshold or as a family grouping. As a threshold, it imagines that harms can be arranged on a scale from very minor to the most major. At some point on this scale, harms become sufficiently major to allow the possibility that they may justify the use of lethal force in order to avoid them. In this, it does not change the nature of BCM but rather simply moves the threshold. The criticisms of BCM that are outlined below might be adapted to any threshold conceptualisation.

Alternatively, vital interests might be regarded as those that allow the possibility of use of lethal force for some other reason, most likely that they share a characteristic that might bestow this possibility. Arbitrarily designating them as centrally important is not yet a principled argument for sufficient reason. And it remains unclear what a principled designation of vital interests is in this schema or, indeed, what the other reason should be. For this approach to work, a major problem would need to be surmounted. That problem is that an explanation is owed as to why certain types of harm are relevant and others are not.

Lazar's point that we might devalue the moral seriousness of war by denying that killing is an intrinsic part of war is flawed. I am sure that he intends that it is taken as a warning not to devalue the moral seriousness of war. However, there is an aspect of this statement that implies that the moral seriousness only emerges from the killing involved. This is false for anyone who subscribes to any of the claims other than claim A. He continues to say (2014:17),

> If we could fight wars in which all those whom we killed were culpable for the threats that we seek to avert, then warfare would not seem such a dreadful thing. Although the good guys will undoubtedly suffer losses too, they can be sure that they will kill only the bad guys, so although there are prudential risks in war, there are no, or few moral risks. It just seems wildly unrealistic to imagine that warfare could be so morally congenial.

Lazar's point is that believing that killing is not essential to war may not represent fully the harm of war because killing is the most important harm in warfare. However, if one accepts that other forms of harm are potentially equally important then it is clear that suggesting that killing is an essential characteristic of war can hide the true harms of a conflict - which clearly is not Lazar's aim. The arguments in this project wholeheartedly support the idea that any large-scale conflict must not be regarded as

‘congenial’ or ‘ideal.’ The difference between this project and Lazar’s point is that it is not only the presence of killing that removes the possibility of congenial war, or any other idealisation of war.

Here, it is worth briefly revisiting the claims that might be made regarding the importance of killing in large-scale conflict if one is to challenge BCM.

A. Killing is the only important moral factor
B. Killing is the most important moral factor
C. Killing is one of the important moral factors
D. Killing is not one of the most important moral factors
E. Killing is not a moral factor

The claim that even a defensive war cannot be undertaken without killing does not directly affect the position one might adopt. Whether or not one believes that killing is a necessary element of war does not necessarily define its importance. The critic of BCM does not need to embrace claim (D) or (E). The point need not be that killing is not important. Such a claim would be ridiculous. The critic only requires a demonstration that (A) is untenable.

The purpose of this section, however, was to demonstrate that body count morality is prevalent across the board. It is exhibited in both quantitative and conceptual analyses of self-defence and large-scale conflict. There are few areas of discussion surrounding self-defence, conflict and war that are immune to its use. In general, it appears that it is accepted as an abstraction or a heuristic device. When the threshold nature of body count morality becomes evidently problematic, theorists tend to retreat to a tiered system, usually involving ‘vital interests,’ in which many of the problems of body count morality persist.

2.4 The Body Count Method

Body count methodology reduces the consideration of the effects and outcomes of conflict to an accounting of participants' deaths. The tendency to express body count morality is displayed across the literature surrounding large-scale conflict. Two instances of this tendency show the range of application of BCM. Firstly, in many quantitative analyses of war, the definition of war itself is grounded in a body count. Secondly, in less quantitative analysis, much of a specific form of philosophical investigation relies on abstract examples in which the complexity of the conflict is reduced in the hope of revealing underlying truths and in which body count morality is used to provide that simplification or abstraction.

These examples aim to show that the use of body count morality is not limited to a particular area of the literature on war or any particular stance. For instance, there is

an analogous division between those who favour abstract examples and those who base their arguments on historical examples. As Seth Lazar (2017:39) says in describing this,

> What kinds of cases should test our principles in the ethics of war? We can think of realistic scenarios, paying attention to international affairs and military history. Or, more clinically, we can construct hypothetical cases to isolate variables and test their independent impact on our judgements.

That such a divide exists is uncontentious. The examples of this section will be on opposing sides of this divide. Other divides exist in the philosophy of war. This section aims to provide a convincing argument that body count morality exists on all sides of these various divides. The use of evidence from quantitative analysis and evidence from the use of theoretical examples demonstrates that the use of body count morality is widespread, pervasive, and largely agnostic of other divisions.

The aim of the Correlates of War (CoW) project is to provide accurate and reliable data for use in the study of war. The description from the CoW website (Correlates of War 2022*)* is instructional regarding the scope and goals of the project.

> The Correlates of War Project was founded in 1963 by J. David Singer, a political scientist at the University of Michigan. The original and continuing goal of the project has been the systematic accumulation of scientific knowledge about war. Joined by historian Melvin Small, the project began its work by assembling a more accurate data set on the incidence and extent of inter-state and extra-systemic war in the post-Napoleonic period. To do this scientifically Singer and Small found they needed to operationally resolve a number of difficult issues such as what is a “state” and what precisely is a “war.” Building upon the work of other pioneers such as Pitirim Sorokin, Lewis Frye Richardson, and Quincy Wright, Singer and Small published The Wages of War in 1972, a work that established a standard definition of war that has guided the research of hundreds of scholars since its publication.

In providing a standard definition of war the project has been at least partially successful. Inevitably there is further discussion surrounding the precise demarcation of war, particularly with the changing nature of war in areas such as cyber conflict, soft war, and hybrid war, but the CoW definition has provided a bedrock for discussions in war studies across disciplines. The definition is based in part on the demarcation of war as a conflict that involves in excess of 1000 fatalities in a year. “We consider,” stated the founders of the CoW project (Small & Singer 1982), “the taking of human life the primary and dominant characteristic of war.”

The prevalence of the CoW project as a data source for war studies has resulted in the fact that a body count methodology is embedded in the resultant quantitative analysis. This may be implicit or explicit but the definition of the context of study based on the number of deaths creates a bias in the study of war.

One might say that the separate consideration of other forms of harm is not excluded by the definition of war by 1000 deaths. But this is entirely too quick. The modern studies of gender and race have shown that the way that definitions are constructed are influential in the discussion. Definitional biases are pernicious. The exclusion from war of actions or situations that do not reach 1000 deaths, even though they may involve the imposition of widespread and significant harms, places them in some other category and removes them from analysis in the context of war studies. As is argued later in this section, that removal results in an inadequate representation of the harms created by large-scale conflict.

If those theorists who tend towards quantitative analysis show bias towards body count methodology, those who tend towards the use of abstract or hypothetical examples fare no better. A fundamental part of the technique of philosophical examples of this type is abstraction, where abstraction is defined as a technique that attempts to remove extraneous information with the aim of reducing complexity. For once, it is worth quoting Wikipedia (Wikipedia 2022a) when it says, "Conceptual abstractions may be formed by filtering the information content of a concept or an observable phenomenon, selecting only the aspects which are relevant for a particular subjectively valued purpose."

In philosophical examples, particularly those which ground a theory of war on a theory of personal self-defence, it is appealing to reduce a complex situation to one in which the only quanta in calculations are single lives. For instance, ‘trolley problems’ involve a moral decision in which one must decide in a variety of circumstances whether an action is permissible that saves ***X*** lives but costs ***Y*** other lives - where both ***X*** and ***Y*** are whole numbers. Philippa Foot (1967:2) was one of the philosophers who brought this type of example into mainstream philosophical thought.

> Suppose that a judge or magistrate is faced with rioters demanding that a culprit be found for a certain crime and threatening otherwise to take their own bloody revenge on a particular section of the community. The real culprit being unknown, the judge sees himself as able to prevent the bloodshed only by framing some innocent person and having him executed. Beside this example is placed another in which a pilot whose airplane is about to crash is deciding whether to steer from a more to a less inhabited area. To make the parallel as close as possible it may rather be supposed that he is the driver of a runaway tram which he can only steer from one narrow track on to another; five men are working on one track and one man on the other; anyone on the track he enters is bound to be killed. In the case of the riots the mob have five hostages, so that in both examples the exchange is supposed to be one man's life for the lives of five.

Even though the conclusions that emerge from trolley problems are often disparate, the actual arithmetic is made simple by the use of whole numbers. However, we can just as easily imagine examples in which, for example, the value of one person’s leg was balanced against the value of three people’s fingers. Such a framing opens the discussion to concerns about the relative value of various limbs and in so doing only

muddies the water by making the arithmetic less clear. Likewise, one can imagine cases that are less cut and dry; cases in which the calculation is more probabilistic than deterministic. For instance, in a typical philosophical example, a surgeon might feel that if she does not remove a limb there is a 50% chance of death, whereas removing the limb obviously has a 100% chance of loss of limb but only a 10% chance of death. Probabilities can also muddy the water. If one's aim is to create a clear example in which the core issue is demonstrated, it makes sense to remove all extraneous details. So, the use of deaths in BCM allows a simple accounting of the relative values of certain options and avoids many of the complications that result from including a wider selection of harms and damages.

The extreme depth of those complications described in the previous paragraph is investigated by Pat Tomlin (2018:13) in a series of examples with titles such as "Finger or Accidental Death,"

> Finger or Accidental Death: Attacker threatens Victim with a broken finger. Victim responds, intending to shoot into the air in order to frighten Attacker, thus distracting him so that he will break Victim's little finger rather than her index finger. She is almost certain to succeed in doing this. However, Victim knows that there is a very small chance that she will accidentally kill Attacker. This will prevent the attack altogether.

In a string of similar examples, Tomlin investigates what he calls the 'separation assumption' - that the goods of an act are to be totalled in isolation to the harms of an act - by arguing that the assumption "rules out taking account of what pairings of harm and good may come into existence together." Tomlin, in what I think is a remarkable paper, engages with both the subjective nature of proportionality calculations and the probabilistic nature of those calculations. However, what is required here is not full engagement with his arguments but an understanding that the complications are indeed complicated, and once one steps away from the solid ground of body count methodology things get extraordinarily complex very quickly. Body count morality is an abstraction that attempts to avoid these complications.

The advantage of relying on a body count methodology is that it eliminates much of the complication and allows focus to be drawn to specific elements of the situation. It is certainly not my aim to devalue abstraction entirely. This type of abstraction can be valuable and is a successful and proven technique in science, philosophy and other disciplines but relies on the reintegration of the conclusions from the abstraction with the unabstracted situation.

The disadvantage of such abstraction is that without reintegration the examples can become divorced from the complex situation on which they are based. In dealing with what Michael Neu (2013:461) describes as "neat fictitious worlds, rather than the complex real world," these abstractions ignore and devalue harms that do not fit into the pristine structure of clear-cut theories.

For body count morality to be a justifiable approach the advantages need to be great enough to justify the disadvantages. If one's relationship with large-scale conflict is purely academic, then it might be that the simplification of the moral arithmetic justifies the disadvantages. From any other standpoint, the disadvantages heavily outweigh the advantages. Returning to the words of the Wikipedia quote, that abstraction works by "selecting only the aspects which are relevant for a particular subjectively valued purpose," it is clear that the process is not entirely objective. If the purpose is to represent the total harms of large-scale conflict, then body count morality does not select the relevant aspects.

Previously, the possible claims regarding the importance of killing were identified.

A. Killing is the only important moral factor
B. Killing is the most important moral factor
C. Killing is one of the important moral factors
D. Killing is not one of the most important moral factors
E. Killing is not a moral factor

The conclusion of this section is that (A) is untenable. If we aim to represent the total harms of large-scale conflict, then we must reduce the claim to at least claim (B) because killing is not the only important factor.

The difference between (A) and (B) is that (B) allows other significant moral factors. Defending (B) might be coherent on the individual level. One might believe that the worst moral harm that could be inflicted on one is death - though even that claim is contested in this project.

Once one has allowed that there are other significant factors, it becomes more difficult to defend the fact that killing remains the most important factor. Put a different way, once one accepts (B), the slide to (C) is likely. This is because it is trivial to construct abstract examples that demonstrate that the protection of a single life cannot justify extensive and widespread harm. At their most extreme this type of example asks such questions as whether we should protect a single individual at the cost of maiming all the rest of the world's population. There may be no easy answer to these questions, but it is not immediately clear that the death of the individual cannot be an ethical answer.

Body count morality's last hope of salvation is that it still provides an adequate representation of the overall level of harm, or that it in some way tracks overall harm - that in some way, although it is not the only metric of harm, it still provides a route by which overall harm may be assessed. It is worth taking the next section to see if this salvation is plausible.

2.5 But it is just a heuristic device, right?

At this point, some readers will be shaking their heads. ‘But’ those people will claim, ‘we do not think that killing is all that is important. It just provides a way of representing harm.’ As Seth Lazar (2020) says,

> This focus on killing might seem myopic—war involves much more violence and destruction than the killing alone. However, typically this is just a heuristic device; since we typically think of killing as the most presumptively wrongful kind of harm, whatever arguments one identifies that justify killing are likely also to justify lesser wrongs.

That seems credible only if certain commitments regarding this heuristic device are met. The first is that there is a proven correlation between killing and other harms. The second is that it assumes that killing is the greatest harm.

The type of correlation that would be required would either be quantitative or conceptual.

A quantitative correlation would suggest that while killing did not represent the entire harm of a conflict it was in some way proportionate to the whole harm, or at least represented the overall harm in an adequate manner. In other words, a conflict that created 100 deaths was twice as bad as one that created 50 deaths. In practical terms, this is obviously untrue. 50 deaths and the long-term effects of radiation might balance differently to 100 deaths from drowning with few long-term effects. Who knows without much greater detail? Perhaps the quantitative correlation is loose but still meaningful? It is like a rule of thumb - ‘the more killing, the worse the harms of the conflict.’ To me, even this seems implausible. Apart from anything else, it depends on the idea that large-scale conflict inevitably involves killing. While this is largely an accepted truth about war, there are persuasive arguments that it is not necessarily true about war (for example, Steinhoff 2009). One of the reasons to talk in terms of conflict rather than war is to avoid these presuppositions. In terms of large-scale conflict, in the context of the changing nature of war, it will be argued that killing is not a necessary characteristic of this type of conflict. If we surmise that some conflicts will not involve killing at all, but still produce high levels of harm across populations, the quantitative correlation seems untenable. But one need not go that far because it is intuitive, as is discussed more fully later, that an outcome with little killing but a huge amount of other harm might not be preferable to an outcome with slightly greater killing but much less other harm.

A conceptual correlation would need to claim that all types of harm were in some way conceptually related to killing. Or, to put it another way, that there were not types of harm that were simply unrelated to killing. This, however, is not tenable. Historically, spiritual harms have been unrelated to physical harms to an extent that would

challenge this conceptual correlation. Physical harm and spiritual harm were so different that one could not be correlated to the other. I will argue that the effects of the digital revolution are such that our digital interests also challenge this correlation. That is left for the next chapter, but even without that support, the idea that there are no harms that do not correlate to killing seems untenable. A series of actions that negatively impacted on long-term women's human rights in a country, but created little killing is the type of example that can easily be constructed to challenge the conceptual correlation. Many types of legitimate harm to populations do not equate to physical harms.

Moreover, there is perhaps a reliance on the quantitative correlation in the conceptual correlation. The two supposed correlations are not unrelated. This is because of the presupposition that physical harm is on a scale that ranges from minor harm such as a bruise to the most major of harms, death. The scale may not be linear in its specifics but, in general, it represents a coherent spectrum of harm from the negligible to the extreme. In such an environment a quantitative correlation might make sense. If rather than a scale, one has an environment in which there are multiple scales or even a dispersed field of harm the idea of a simple correlation is more difficult to sustain. If one believes that spiritual welfare is of significant importance this broadens and diffuses the linear scale. Likewise, if one believes that political freedom is of significant importance that broadens and diffuses the linear scale. In terms of this project, if one believes that the digital revolution has altered those things that are most important to us, then that broadens and diffuses the linear scale. In this environment, the idea of a quantitative correlation between killing and harm seems difficult to sustain.

The second commitment that must be accepted if killing can be used as a heuristic device as described is that killing is the greater, or greatest, wrong. Lazar claims that the logic is that "whatever arguments one identifies that justify killing are likely also to justify lesser wrongs." A critic of his might ask whether they also justify the greater wrongs. This flippant response disguises a serious challenge; that killing is not the greatest wrong.

Susanne Sreedar dedicates a section in her paper (2008:788-789) about Hobbes' attitudes regarding this issue to discussing the "Death as Worst Possible Evil" premise. Her point is that, although Hobbes is widely credited with thinking that death was the worst evil, he, in fact, acknowledges that there are cases where it is not.

> Hobbes recognizes that people might rationally prefer death to life if their lives are sufficiently painful (De Homine, 11.6). Second, he acknowledges that "there are commands that I would rather be killed than perform."

The idea that death is not the worst of all evils in all cases is defended more fully later but its truth can be sketched both on the individual level and the political level. If killing is the greatest harm, why would a parent conceivably choose their own death rather than a lesser harm for their child? If killing is the worst harm, is the phrase 'a fate

worse than death' meaningless? If killing is the greatest harm, the idea that some causes beyond saving lives are worth dying for that motivates at least some instances of war is also untenable.

As both the first and second commitments cannot be met, the use of killing as a heuristic device is inappropriate. It makes the analysis clearer and removes conceptual messiness, but in doing it distances the analysis from the realities of large-scale conflict.

Finally, there is a linguistic argument against the use of 'killing' as a heuristic device. That is that its use subtly, and not so subtly, skews the conversation. It removes legitimate consideration of other harms that affect the lives of individuals and populations. In an odd twist, the use of 'killing' sanitises the conversation in that the discussion is quantified in whole numbers rather than accepting the messy and analogue horrors that truly represent the effects of large-scale conflict. A war that is quantified by a certain number of deaths is perhaps more palatable than the same war if one considers deaths, maimings, rapes, psychological damage to both participants and others, and the myriad other factors that make up a large-scale conflict.

2.6 Harm is Not Fully Represented by Body Count

In an important 2005 paper, Jeff McMahan (2005:388) uses the example of the tactical bomber,

> **Tactical Bomber**
>
> A tactical bomber fighting in a just war has been ordered to bomb a military facility located on the border of the enemy country. He knows that if he bombs the factory, the explosion will kill innocent civilians living just across the border in a neutral country. But this would be a side effect of his action and would be proportionate to the contribution that the destruction of the facility would make to the achievement of the just cause. As he approaches, the civilians learn of his mission. They cannot flee in time but they have access to an anti-aircraft gun.

Apart from the slight concerns that these are strange civilians to have a handy anti-aircraft gun (but then people in philosophical examples do appear to have access to extensive weaponry), this example has been used in various forms to elucidate the influence of intention on collateral killing. I am going to simplify and repurpose the example.

> **Tactical Hacker**

> Tactical hackers involved in a large-scale conflict have been ordered to disable a military facility located on the border of the enemy country. They have discovered a vector by which they can shut the facility's operations down. They are confident that the action will not directly cause any fatalities, but it does involve disrupting the power supply over a large area, including the supply to major sanitation and water supply plants.

In both examples, there is an implicit value in the disabling or destruction of the military facility. The benefit of doing so is not detailed but is assumed to be sufficient to justify the harms created. How this justification is structured is at this point left open, though it is likely to be an application of proportionality (which is discussed more fully in Chapter 5). What must be assumed is that there are at least certain circumstances in which the actions may be regarded as justified.

The problem that the Tactical Hacker highlights is simple. If one uses BCM and if the actions of the hackers do not directly cause any deaths, the proportionality assessment is rendered useless. On one side of the assessment is a benefit and on the other side of the assessment are the deaths - of which there are none. In body count terms, the action of the tactical hackers will inevitably be proportionate because a body count cost to their actions does not exist. This seems wrong because it allows the imposition of a range of significant harms as long as no killing is involved.

Rather, it seems entirely plausible that the aggregation of widespread harm might well result in a circumstance where the hackers' successful prevention of the future harm of the military facility was disproportionate because the aggregated harm outweighed the value of the destruction/disabling of the military base. And this assessment would be valid even if there are no deaths on the side of the balance associated with the preventative action and one or more deaths on the side of the prevented harm. The issues and subtleties of this kind of proportionality calculation are the subject of Chapter 5 and so will be left unresolved here. The point that is critical for the argument is that if harm is only assessed in terms of deaths, one might reach an incorrect assessment of what actions are justified.

At this stage, I am not suggesting that the understanding of those who use BCM is as coarse or unsophisticated as the calculations being detailed here. The argument in the previous paragraph is largely uncontentious in this respect, and no reasonable theorists are likely to dig their heels in and claim that we must only tally the deaths in such a circumstance. The aim is to provide an extreme example of how the use of BCM can result in inappropriate ethical assessments. If this occurs in extreme examples, it is plausible that a similar effect is seen in less extreme examples.

It would be easy for an advocate of theoretical examples to design cases which display that there are other significant harms involved in large-scale conflict and that those harms either individually or aggregated across populations are as significant as killing. How might we compare a war waged only with non-lethal, but disabling and

disfiguring, chemical weapons to a war waged with conventional weapons? Can we not imagine lurid examples that demonstrate that there are indeed fates worse than death? Or examples aimed at emphasising the sentiment that it would be kinder simply to kill than behave in other violent, degrading, or coercive manners? Examples can be made increasingly lurid and sensationalist until even the greatest sceptic must acknowledge that other harms are comparable to killing - or even worse than simple killing. Here, abstract and theoretical examples do not help us in that they become increasingly fictionalised and, in order to discern whether a body count methodology adequately represents the harms of war, we inevitably must turn to real-world examples.

Since 2011, Syria has been in a state of civil war. In 2019, Encyclopaedia Britannica designated this conflict as the second most deadly conflict of the 21st Century (Ray n.d.). The total death toll of the conflict is hard to quantify accurately but is widely believed to be in excess of 350,000 (UN News 2021) many of whom were civilians (Statista Research Department 2021). The conflict in Syria is multi-sided and politically complex. Multiple internal forces are involved with the support of various external powers. In every way, including the death toll, the war in Syria meets the CoW requirements to be designated as a war.

While the significance of hundreds of thousands of deaths cannot be overstated, anyone who has followed the reporting on the conflict must find it difficult to believe that these deaths fully represent the significant harms that have been created by the conflict. A section of the Encyclopaedia Britannica entry demonstrates the political complexity of the situation and the widespread violence that was created,

> Rebel groups seized huge swathes of territory, and the area under government control was reduced to a small strip of land in western Syria. Assad resorted to increasingly desperate and savage measures to maintain power, dropping crude “barrel bombs” on urban populations and using chemical weapons on rebel-controlled territory. (Ray n.d.)

Many of us will recall with horror the images that emerged from the conflict, most of all small children suffering the effects of chemical attack. The Syrian economy was destroyed. GDP per capita has dropped from the region of $60k to approximately $12k (CEIC 2020). As an example, almost a quarter of all housing is estimated to have been damaged or destroyed and this damage alone is valued at $28 billion, roughly twice the overall GDP of the nation. Crime has increased dramatically across the country, and it was estimated that over 6 million people were internally displaced within Syria in 2018 (Statista Research Department 2021). The effect of the conflict has stretched far beyond the Syrian borders. By 2016 there had been over 6 million refugees from Syria, and the flood of refugees had strained the resources and political situations in Europe and other regions.

This is a simple and cursory snapshot of the history of the conflict in Syria, but it is enough to cast doubt on the standpoint that the hundreds of thousands of deaths are truly representative of the amount of overall harm that the conflict has caused. Claiming that the number of deaths, or the overall body count, accurately represents the nature of a conflict is surely a reduction that is only possible in safe confines, isolated from the actuality of war itself. Large-scale conflict is horrific not only because of the deaths involved but also because of the immediate and long-term adverse effects on populations with regard to health, wellbeing, economic prosperity, personal freedoms, political justice, the mental health of participants and diverse other areas. It is hard to imagine any area of life which is immune from the horrors of war. The effects extend beyond the spatial and temporal limits of the conflict and are not limited to those who die.

A. Killing is the only important moral factor
B. Killing is the most important moral factor
C. Killing is one of the important moral factors
D. Killing is not one of the most important moral factors
E. Killing is not a moral factor

In terms of the list of possible claims that was presented previously, the example of Syria certainly supports claim (C). It also denies claim (A). The status of claim (B) in the Syrian context is open to argument. This supports the conceptual argument that was presented earlier in which it was concluded that killing was not necessarily the most important moral factor, although in certain cases it might be.

2.7 Why We Should Avoid Body Count Morality

Body count morality is widely used in discussions surrounding large-scale conflict because it provides an easy way of quantifying the harm of conflicts. The problem is that it does not adequately reflect the harms of conflicts, particularly in the modern world. This is true in conventional warfare, as examples such as the conflict in Syria demonstrate. However, in the context of the changing nature of conflict, it is increasingly true.

The digital revolution has changed both the methods in which we can inflict harm and the things that we value. It is undeniable that we value our biological lives. However, historically it has always been understood that we also value - and at times more - our spiritual lives, our political lives, and our intellectual lives. The digital revolution has added one more life to that list; we value our informational lives. That fact means that body count morality is even less appropriate now than it ever has been.

In that context, cyber methods and cyber interests provide perhaps the strongest motivation for avoiding body count morality. There is a developing theme that is

discussed in Chapter 4 that Cyberwar may form an ideal form of warfare in that it will achieve strategic aims without causing deaths (e.g., Jenkins 2016). This is a dangerous road to travel down as it denies the significant harms that are created by this 'ideal' form of conflict and allows nations and other actors to inflict harm in what may become a deregulated environment. To assess fully the harm created in large-scale conflict we must avoid body count morality. As Michael Neu says when discussing the tragedy of war (2013:466), "killing, of course, not being the only morally relevant harm that can be caused, or avoided, by waging or not waging war." We must be sensitive to all the morally relevant harms rather than just killing.

3. Cyber Methods and Cyber Interests

All things physical are information-theoretic in origin and this is a participatory universe ... Observer-participancy gives rise to information.

John Archibald Wheeler

3.1 Introduction

The previous chapter argued that the dominant way of assessing the harm of conflict, which was termed body count morality, was not appropriate overall, and that it was particularly inappropriate in the context of cyber conflict. A more appropriate technique or process for evaluating harm is required. Harm has its own definitional challenges. In the broadest of terms, the approach that is adopted in this chapter is that harm equates to the interference with, or degradation of, interests.

This allows us to pivot from talking in terms of harm to talking in terms of interests. We should be wary of forgetting that this pivot has been made. There are other ways in which harm might be conceptualised. The adoption of a broad interpretation of harm is intended to generalise the conclusions that are reached but it will be necessary to assure ourselves that this route has not limited or devalued those conclusions.

The emphasis on interests elucidates a key fact. Our interests are plural, and this forms the core theme of this chapter. Our framework for assessing harm must be pluralistic because our interests are pluralistic. Various conceptions can be united by the understanding that there are a group of factors that contribute to the achievement of our interests. Rodin talks in terms of vital interests. Rawls talks in terms of basic liberties and primary social goods. Nussbaum talks in terms of capabilities. In a broad, yet meaningful way, all these conceptions (and many others) posit a plurality of factors that contribute to a fulfilled and valuable human life.

One of the axioms of this project is that *We value Flourishing Human Lives,* and it is at this point in the project that the importance of this axiom is most telling. A flourishing human life is the quanta of assessment that is used in this project. That the factors that contribute to a flourishing and valuable human life are plural is widely accepted although the language used to describe this pluralism varies widely. A choice is made in this project to use language associated with objective list theories of welfare because these theories reflect an inherent pluralism - although it is worth noting that other theories of welfare such as hedonism and desire theories do not necessarily exclude pluralism of human interests. It is a choice to use the language of objective lists, and the argument could equally be expressed using other language.

Section 3.2 outlines this pluralism and uses one application of an objective list, the capabilities approach, to gain insight into how the pluralism might be applied and what general observations can be drawn from this specific application. Sceptics of the capabilities approach need not be discouraged. The project only assumes that some form of pluralism is plausible.

The following two sections (3.3 & 3.4) demonstrate that informational factors not only can be included in that pluralism, but must be included. In the first of the sections in this block, the importance of informational factors in our lives is emphasised. The idea of informational interests is developed, in particular the categories of private data control and public data access. The second section fleshes out this conception of cyber interests and outlines two prospective interests, private data control and public data access. These interests will be developed and described as this chapter - and the rest of the project - continues.

In the subsequent section (3.5), the unique ability of cyber methods to influence these interests is described. This provides the answer to the question of whether we should be concerned with cyber methods and how much we should be concerned. The answer to the first question is that we should and the answer to the second question is that cyber methods have a unique ability to influence both our informational and physical interests. Although their effect on our physical interests has dominated analysis so far, that may not be truly representative of the capabilities of cyber methods. In fact, the informational effects of cyber methods are equally, if not more, significant.

Finally, in section 3.6, a pair of strong objections to this line are rebutted. The first is that, despite talk of informational factors, the biological body is of primary importance. In answer to this, it is argued that a biological body is necessary for a flourishing human life but not sufficient. The second is that the talk of the importance of informational aspects of our lives is a ‘first world’ phenomenon. In answer to this, two observations are made; that information is critical to flourishing lives even in cases where the individuals are unaware of that information and that as the digital revolution progresses one would expect the phenomenon to become truly global.

The previous chapter rejected the most prevalent monist interpretation of assessing the value of human life, body count morality. This chapter offers a pluralistic alternative that includes recognition of the importance of informational interests. Because those informational interests are critically related to human lives and because digital technologies are the dominant form of informational technology, those interests may be termed informational interests, digital interests, or cyber interests. The chapter starts by talking of informational interests and ends in talking of cyber interests. As is explained, this is a semantic difference rather than a substantive one. What emerges is that there are two potential families of informational interests that need to be incorporated in the plurality of factors that contribute to a flourishing human life. These are private data control and public data access.

3.2 Pluralism of Human Interests

This subsection aims to provide a plausible form of pluralism on which subsequent arguments can be based. At first glance, this may appear tangential to a discussion of cyber conflict. However, large-scale conflict can be regarded as the institutionalised imposition of harm for political motives. For that process to be effective in reaching a political goal, the harm must be aimed at human interests that are significant to human lives. In the previous chapter, I argued that these interests could not be represented by biological life alone. The significant harms of war are diverse. As Michael Neu (2013:466) says when discussing the tragedy of war, “killing, of course, not being the only morally relevant harm that can be caused, or avoided, by waging or not waging war.”

Although body count morality is dominant, the importance of other factors than biological life is recognised in the philosophy of war. A pertinent demonstration of this is the case of the 'bloodless invasion' that was discussed in the previous chapter which implicitly accepts that there may be other things of value than body count. Saba Bazargan-Forward (2017:145) discusses these issues when he says,

> The bloodless invasion dilemma assumes that the political rights the invaders violate are non-vital, in that they are not the sorts of rights that we can defend at the cost of innocent lives. But bloodless invasions typically violate not just political rights but our interests in retaining the capability of living recognizably meaningful lives. And these are indeed vital – or so I will argue.

It is the phrase ‘our interests in retaining the capability of living recognizably meaningful lives’ that is pertinent to the following discussion. The use of the word ‘interests’ reflects an inherent plurality. The use of the word ‘capability’ echoes the influential capabilities approach that has transformed the study of development and freedom in the last 30 years which will be used as an example in this section. The idea of a ‘good life’ or a ‘meaningful life’ or a ‘truly human life’ is a typical value that is promoted in writing supporting a capabilities approach that emphasises that not only just being alive is important in people’s lives. It reflects one of this project's axioms; that the correct arena of social and political theory is the development of flourishing human lives.

As outlined in the introduction to this chapter, we move from talking in terms of harm to talking in terms of interests. This collection of human interests may also be regarded as human well-being. The conceptual assumption that allows this is the idea that harm equates to the degradation of interests and the degradation of well-being. This is one of the axioms of the project and allows us to think in terms of the degradation of human interests as a representation of harm.

There are usually considered to be three types of philosophical theories of well-being: objective list theories, desire theories and hedonism.[17] Hedonism assesses well-being in terms of the balance between pleasure and pain. An act or event that increases pleasure or decreases pain increases well-being. Desire theories assess good in reference to the fulfilment of individual desires. An act or event that fulfils a desire increases well-being. Objective list theories provide a list of the interests that contribute to well-being. An act or event that causes an increase in any of the listed interests is good.

None of the three types of theory of well-being exclude a plurality of interests that contribute to well-being. A basic version of hedonism might suggest that only happiness mattered, but it would be impossible to say that there was not a plurality of factors that contributed to that happiness. Desire theories suggest a plurality of desires. It seems implausible that our desires might be anything other than plural. Objective list theories are inherently pluralistic and for this reason, much of the remaining discussion in this project is framed in terms of well-being represented in objective list terms.

The remainder of this section looks at a particular way in which an objective list theory can be realised. The capabilities approach has its foundation in the work of Nobel laureate Amartya Sen. It has been influential in governmental and UN policy and has been widely championed and developed, most notably in the philosophical sphere by Martha Nussbaum. In *Sex and Social Justice*, Nussbaum (2000:39) asks the question "What activities characteristically performed by human beings are so central that they seem definitive of a life that is truly human?" The sophisticated and complex literature that has developed around the capabilities approach is grounded in the concept that in the consideration of welfare both for individuals and populations what is important is not the resources available to individuals but the overall potential of individuals to use those resources. A capability, in this framework, is the freedom and ability to implement valuable outcomes.

Nussbaum answers her question by providing a list of ten capabilities that are essential for a meaningful or truly human life. Nussbaum's particular list is both widely used and widely contentious. It may be true that any such list is likely to be contentious since it almost by definition claims universality for itself in ways that may be problematic (e.g., Jaggar 2006 and Robeyns 2016). Certainly, Sen has always chosen to avoid a definitive list, and Nussbaum has always claimed that her list of capabilities is open to revision by further discussion and analysis. Nonetheless, Nussbaum's list in *Sex and Social Justice* (2000:41-42) provides a widely discussed objective list theory that provides a useful example of how such a theory might be fleshed out.

---

[17] For the standard interpretation, see Crisp 2021. For examples of challenges to this, see Woodard 2013.

1. Life. Being able to live to the end of a human life of normal length & not dying prematurely.

2. Bodily health. Being able to have good health, including reproductive health; being adequately nourished. Being able to have adequate shelter.

3. Bodily integrity. Being able to move freely from place to place. Being able to be secure against violent assault, including sexual assault. Having opportunities for sexual satisfaction and for choice in matters of reproduction.

4. Senses, imagination, thought. Being able to use the senses; being able to imagine, to think, and to reason--and to do these things in a way informed and cultivated by an adequate education. Being able to use imagination and thought in connection with experiencing, and producing expressive works and events of one's own choice,.

5. Emotions. Being able to have attachments to things and persons outside ourselves. Being able to love those who love and care for us. Being able to grieve at their absence, to experience longing, gratitude, and justified anger. Not having one's emotional developing blighted by fear or anxiety.

6. Practical reason. Being able to form a conception of the good and to engage in critical reflection about the planning of one's own life. (This entails protection for liberty of conscience.)

7. Affiliation. Being able to live for and in relation to others, to recognize and show concern for other human beings, to engage in various forms of social interaction; being able to imagine the situation of another and to have compassion for that situation; having the capability for both justice and friendship. Being able to be treated as a dignified being whose worth is equal to that of others.

8. Other species. Being able to live with concern for and in relation to animals, plants, and the world of nature.

9. Play. Being able to laugh, to play, to enjoy recreational activities.

10. Control over one's environment. (A) Political: being able to participate effectively in political choices that govern one's life; having the rights of political participation, free speech and freedom of association. (B) Material: being able to hold property (both land and movable goods); having the right to seek employment on an equal basis with others.

After proposing this list Nussbaum immediately clarifies that each of these capabilities is necessary for a meaningful human life and that while each capability is independently important, they are related in complex manners. An extended quotation (2000:42) is valuable,

> The "capabilities approach," as I conceive it, claims that a life that lacks any one of these capabilities, no matter what else it has, will fall short of being a good human life. Thus, it would be reasonable to take these things as a focus for concern, in assessing the quality of life in a country and asking about the role of public policy in meeting human needs. The list is certainly general—and this is deliberate, to leave room for plural specification and also for further negotiation. But like (and as a reasonable basis for) a set of constitutional guarantees, it offers real guidance to policymakers, and far more accurate guidance than that offered by the focus on utility, or even on resources. The list is, emphatically, a list of separate components. We cannot satisfy the need for one of them by giving a larger amount of another one. All are of central importance and all are distinct in quality. This limits the trade-offs that it will be reasonable to make and thus limits the applicability of quantitative cost benefit analysis. At the same time, the items on the list are related to one another in many complex ways.

The extended list of capabilities gives the approach an inherently pluralistic nature. Furthermore, the emphasis that the capabilities are independent of each other results in the understanding that an excess of one capability cannot entirely compensate for a lack of another. Nussbaum's point is that what gives life moral weight is not just biological life itself but a meaningful and truly human life. It is by the fact that "human capabilities exert a moral claim" (Nussbaum 2000:42) that human lives have moral value. The extension of this is that when the capabilities are not present in an individual's life, that life is undermined in important ways.

> When these capabilities are deprived of the nourishment that would transform them into the high-level capabilities that figure on my list, they are fruitless, cut off, in some way but a shadow of themselves. They are like actors who never get to go on the stage, or a person who sleeps all through life, or a musical score that is never performed. (Nussbaum 2000:43)

On the political level, the value of human lives is framed in terms of the development of basic human capabilities and that provides the basis for the duties of government.

> We believe that certain basic and central human endowments have a claim to be assisted in developing, and exert that claim on others, and especially, as Aristotle saw, on government. Without some such notion of the basic worth of human capacities, we have a hard time arguing for women's equality and for basic human rights. (Nussbaum 2000:43)

As described above, Nussbaum's grounding of human moral value in this Aristotelian, or universalist, way has been widely challenged. What remains key in the capabilities approach is that human capabilities provide a metric of welfare and of a flourishing life. Capabilities are formulated in an inherently pluralistic way and so a flourishing life must be assessed in an inherently pluralistic way.

The capabilities approach has transformed the landscape of analysis of human rights, development, and freedom in the last 30 years. There are two major parts to that transformation. Firstly, the use of a single metric to evaluate welfare is ineffective and prone to neglecting individuals and sections of populations. Moreover, this fact can be abused by unscrupulous governments. Secondly, a conceptual movement away from simply thinking that life is enough and instead demanding meaningful, and truly human, life has allowed system inequalities and injustices to be challenged.

It is surprising that scant attention has been given to applying a parallel methodology to conflict and war.[18] As has been argued in the previous chapter, it is quite clear that a simple body count is not representative of the significant harm created by large-scale conflict. Any quantification of harm that ignores entire categories of harm is flawed, and the idea that no group of harms can be as significant as killing is pernicious because it allows the neglect, often willful neglect, of significant harms. In contrast, if one finds the underlying messages of the capabilities approach persuasive in terms of welfare - that human value is based on multivariate capabilities - then it seems more than plausible that harm should be represented by a degradation of those capabilities.

Nussbaum's point, that "a life that lacks any one of these capabilities, no matter what else it has, will fall short of being a good human life" is simply a mirror for the claim that any action that causes the degradation of any one of these capabilities is damaging to a good human life and as such can be morally weighty. Harm in large-scale conflict cannot be evaluated simply by a body count. Rather, it must be evaluated in a pluralistic, multivariate fashion that acknowledges the significance of harms other than killing.

This section has focused on the structure proposed by Nussbaum. That is due to the fact that her structure is most widely adopted rather than any explicit support for her particular list of capabilities. It was presented as an example of how an objective list theory might be effective at describing the plurality of interests that contribute to a flourishing human life. Two important points remain to be made regarding the capabilities approach, both of which are relevant to all objective lists.

The first is that there is a question regarding how fine-grained an objective list can or should be. For instance, it seems less contentious to claim that the family of 'emotions' should appear in a list of capabilities than to itemise which elements should appear in that family. Most would agree that emotional wellbeing is constituent of a flourishing life, defining what elements of 'emotions' should define that family of concerns is more problematic. For this reason, in the following discussion, consideration is primarily given to broad families of cyber capabilities while the exact elements in those families remain to be defined.

---

[18] The only exception of which I am aware is Baker& Roberts 2007

The second relevant point concerns the dynamic nature of any objective list. Sen (2004:77) has notably resisted publishing any version of a list of capabilities.

> The problem is not with listing important capabilities, but with insisting on one predetermined canonical list of capabilities, chosen by theorists without any general social discussion or public reasoning. To have such a fixed list, emanating entirely from pure theory, is to deny the possibility of fruitful public participation on what should be included and why.

Most importantly for the current project, Sen is clear (2004:78) that he does not think that any list should freeze the capabilities for "all societies for all time to come, irrespective of what the citizens come to understand and value". It is this concept that empowers theorists to re-evaluate any collection of significant capabilities on the basis of societal change. The digital revolution has been a transformational and widespread societal change that has greatly influenced the range of things that citizens have come to understand and value. Neglecting those changes in any list of capabilities results in a misapprehension of those things that are significant in people's lives.

The capabilities approach gives us an example of how an objective list theory might be developed and used. Its influence on development and welfare has been undeniable. At the core of any objective list is the idea that a multitude of factors contributes to flourishing lives. Nussbaum's list gives one an idea of how such a list might be implemented, but it is also clear that no list is set in stone. The list must be able to evolve along with societal changes.

Transferring the observations from the capabilities approach to a more general interpretation of an objective list allows us to posit that our interests should be assessed in a pluralistic and multivariate fashion and that those interests are likely to change as society changes. In the rest of this project, the word 'interests' will be used in the context of this understanding. It is worth stressing that this understanding does not necessarily rely on an adoption of an objective list view of welfare. Readers who prefer other theories of welfare may still accept the need for a plurality of interests that are required for flourishing.

3.3 An Informational Body

The aim of the subsequent sections is to persuade that informational factors need to be included in the list of factors that contribute to a flourishing human life; those interests linked to our informational lives are significant. A first step in that is to define some language, in this case, the use of the term 'informational body'. That language stems from the thought that humans are informational creatures. What that means in precise terms is less clear. The idea that information and information processing is critical to humanity is central to this idea, but there are disparate ways in which this informational characteristic might be conceptualised. This section aims to navigate a path through this landscape.

Before the main argument is entered a small amount of preparatory work is required. What exactly is an informational creature? Theorists as diverse as Lucian Floridi (2016), Andy Clark (2004) and Rosi Braidotti (2013) have discussed the details of humanity's relationship to and integration with information. For our purposes here, a broader understanding is sufficient in which the focus is simply on the fact that humans have a relationship with information, rather than any metaphysical discussion of the significance of this fact.

There are two distinct systemic traits by which the informational nature of humanity is displayed. The first trait is the fact that we are capable of interacting with information. The information with which we may interact might be the location of the local store, one's mother's phone number or a philosophy paper one is reading. The second trait is that we process information into new forms and in doing so feel some form of ownership over it. This might be a conscious or directed process such as writing a poem in which we arrange information in a particular way for particular effect. It might also be less directed. By simply existing we produce a history; a history of where we were, what we did, and with whom we interacted. This history is informational.

When I think of my - now deceased - father, I think of both the lean athlete who ran competitively in his twenties, and I think of the slightly overweight academic of his eighties. I think of the books he wrote, the students he mentored, but most of all I think of the things he said, the things that he thought. My father, in this purely subjective context, constitutes a bundle of information in my mind. Objectively, as well, his life is defined as much by the information he produced and created, as by what he achieved physically. That was true as much when he was alive as now that he is no longer alive. In this paragraph, there is information that is 'about my father' and information that my father created. Both categories are significant.

In discussing this type of bundled information, a useful terminology is to talk in terms of the 'informational body'. The informational body represents the collection of information that is associated with a human life. For those who struggle with the use of 'body' in this context, it perhaps helps to think of it as a 'body of information' associated with a person, or perhaps more accurately an 'informational identity-forming body of information'. The convoluted nature of that final phrasing explains why the phrase 'informational body' will be used. Matilda Arvidsson (2018:16) says,

> Digital bodies comprise the large set of data we produce as we interact with and through the digital technologies available to us. Although these sets of data are made lethal use of for targeting purposes in high-tech warfare, the data is mostly generated by the uneventful everyday uses of cell phones, email programs, apps, and 'smart' household items such as 'connected' refrigerators.

Arvidsson's definition suits the argument that she makes in the paper but is potentially limited. If the digital body is only that data that we produce, then the definition might exclude data that is produced without our intention but is related to us in some other significant way. An example of that type of information would be information that is

created by other individuals or systems that relates to us. It is not produced directly by us but is related directly to us. A broader definition of the digital body would also include that type of information, and that is the definition that is used in this project. To distinguish, I use the term informational body.

More importantly, there is a significant difference between ‘digital body’ and ‘informational body’. That is simply dictated by the fact that a digital body is created by our interactions with digital technologies. This creates a ‘large set of data’. Although not stated, it is also implied that this set of data is digitally stored. If that is the case, the data set of the digital body is a subset of that data that exists about an individual. Firstly, there is information about that individual that is stored in non-digital ways, and secondly, there is information that arises in manners unrelated to digital technology. Our digital bodies are a subset of our informational bodies.

Does this nit-picking about the difference have any real significance? For Arvidsson’s purposes, perhaps not, and there is no criticism here for the definition that she uses. However, if the aim is a full understanding of the effects of the changes of the digital revolution, then the distinction is important. We have always had an informational body, but the digital revolution has changed both the nature of that informational body and the relationship of that informational body to our physical body. Those fundamental changes are at the heart of the digital revolution.

The nit-picking is significant in another way. Using the term ‘digital body’ inevitably imports interpretations used in other fields. The term is closely associated with, and significant in, the fields of transhumanism and posthumanism (e.g., Harraway 1990, Braidotti 2013, Hildebrandt 2016, Bostrom 2011), and in particular the feminist interpretation of these fields. Those fields deal primarily with the nature of humanity and its relationship and position in the universe. The metaphysical conclusions of those discussions are not necessary here. The required commitment is only that the digital revolution has significantly altered society in general, and specifically has altered those things which are valued by individuals. How that affects what it is to be human is outside the scope of this project.

So, the term ‘informational body’ is used for two reasons. The first is to suggest that humans have always been informational creatures whose lives have been defined by the information they interacted with and created. The second is a hope to avoid the more metaphysical implications of the changes of the digital revolution. Those implications, while fascinating, are not needed in the line of argument to be followed.

As a thought experiment, it is worth examining what an informational body might include prior to the digital revolution. Returning to my father, most of his life was spent prior to the digital revolution. That did not mean that he did not have an informational body. For instance, critical to his academic life were his degree, his work history, his published books of poetry, and his academic books. This is all data that to a certain extent defined his life. In an academic context, if someone wanted to know ‘who he was’ that information would provide the answer. He also was a British citizen working in Canada for most of his life. He had a British passport, a birth certificate and various

paperwork from the Canadian government that allowed him to do so. Again, this information defined his life in practical ways. There is much more information that played that role in various aspects of his life. Removing this information from his life - at any point during his life and even now - would radically alter that life. There is no doubt that he had an informational body even before the digital revolution.

The target of this chapter is the understanding that our informational body is an important constituent part of who we are. When phrased in that manner the understanding that our information has deep import in our lives seems relatively uncontentious. Most people are comfortable accepting that they are users, processors and creators of information and that that information and our relationship with it form an important part of our lives and identities. So, the claim is that the informational nature of humans is innate. As such, this informational nature is not the result of the digital revolution but rather an innate characteristic of humanity.

3.4 Cyber Interests

The previous section outlined a conception of the significance of our informational bodies. This significance must be reflected in the plurality of interests that contribute to a flourishing life. In the same way that aspects of our physical body and its security are reflected in our interests, so must aspects of our informational body and its security.

The digital revolution is transforming the way in which our informational bodies are instantiated and function. ICTs have transformed our informational lives and our informational bodies. It is worth remembering that ICT is an informational and communication technology. The fact that the predominant usage is to talk in terms of digital technologies is only an artefact of the overwhelming success of those technologies at this current point in history. That fact is critical in understanding a theme in the challenges to the arguments presented here. Many people object to the statement that digital aspects of our lives are critical to human flourishing. Fewer people object to the statement that informational aspects of our lives are critical to human flourishing.

The aim of this chapter is to provide a framework of analysis that can be applied to the ethics of cyber warfare. This section provides a summary of the informational interests that exist in our lives because those interests are things that we value. Whether we call them informational interests, digital interests, or cyber interests is immaterial. They are interests that relate to our increasingly developed informational lives and are largely mediated through digital technologies. If we accept that the term ‘cyber’ refers to the relationship between human and informational factors, then these interests are intrinsically cyber interests.

We stand in a transitional period - only at the start of the major changes of the digital revolution, so any list of cyber interests will be tentative, subject to change and anticipatory in ways that might open it up to charges of science fiction. However, even

at this transitional stage, it seems clear that two families of interests are emergent. We cannot yet predict exactly how those families are to be instantiated in the future, but it seems likely that they will feature in some form.

The first is control of one's own information. This will incorporate interests regarding information privacy, information autonomy and information ownership. Already, with that provisional list, the complications and developing difficulties of specifically defining that interest become apparent. However, it seems inevitable that as digital information becomes increasingly important in people's lives, there will be a cluster of significant interests that surround the issue of the control of one's own information. I will call this family private data control.

An example of private data control is the control of medical records concerning the history of an individual. The furore when governments suggest releasing medical information to private companies (e.g., Adams 2013) makes clear that there is a strong belief that medical records should be controlled by the individual to whom they apply. In the UK, the battle regarding the privacy of health records is currently (in June 2021 – see Clark 2021) being fought once again. In 2014, an attempt was made by the government to make health data available to research and commercial organisations which was stymied by public campaigning (Fiveash 2014). Whether the current campaign will prove successful remains to be seen. On one hand, the fact that the struggle has been ongoing for over seven years demonstrates that the data has immense value both in monetary terms and as a functional resource. On the other hand, it demonstrates that there is a strong belief that individuals inherently have rights concerning data that is associated with them.

Of course, one might believe that we only care about our information being released to outside agencies because of what those agencies are going to do with it – that our worry is not with the privacy of the data at all, but rather with the instrumental way in which the information might be used. It is worth noting that in the previous example there is no description of the harms that the agencies receiving this information were intending. It was enough that they were receiving the information. A strong emphasis on privacy in general, and on the Internet in particular, suggests that we care deeply about the control of our private data for both fundamental and instrumental reasons.

This seems like strong evidence that we care about control over our personal data, and that care is not grounded entirely in pragmatic worries that this information might be used against us. Rather, it seems the case that we would care about the unauthorised sale of our data even in cases where we were assured that it would result in no physical or emotional harm to ourselves. The control of our data is a concern that is not only grounded in other concerns.

In the context of developing cyber interests, it will be argued that medical records are paradigmatic in that we have always had medical records. We were happy that such records were kept by our doctors when it was clear that those records were kept in a largely inaccessible format i.e., on paper in illegible handwriting. In the post-digital-revolution world, where those records are stored digitally and uploaded to the cloud,

their security and ownership become more problematic. As is discussed in the following section, it is the persistence and accessibility of these records that leads to vulnerability and that causes us to value control over the data.

Conversely, the second family of interests clusters around access to common data and information. Again, the specifics of this family will develop over the coming years, but issues such as access, sovereignty, informational trust, the ownership of data and the ownership of physical infrastructure will feature. The ability to interact with data that is regarded both as public and reliable will continue to be a critical interest. Interruption, denial, or manipulation of this access can be regarded as coercive and certainly genuinely harmful. I will call this family of interests, public data access.

A salient example of public data access is presented in the discussion that surrounds net neutrality. The idea that underpins net neutrality is that the owners and managers of information networks that are essential to our lives should not influence the information that is transported on that network. For example, the providers of the internet backbone, who are private entities, should not filter, censor, or otherwise influence the information that is transported. Net neutrality is at the heart of the ethos of an open internet but produces problems for large organisations such as nations because the freedom of information transfer can often be used against the interests of those organisations. The discussion is complex.

There is a tension between private data control and public data access that mirrors the issues of property, ownership, freedoms, rights, and sovereignty that exist in the physical world. The two can conflict in various ways. An example of this occurs in the context of criminal records. Should the individual be able to control access to this information or is it in the public interest that it is made available? Does public data access trump private data control in this case, and if so, how do we define the cases where that is acceptable? These questions are not in the scope of this chapter but do illustrate that the two families of interest frame problems clearly and also illustrate that the families of interest can come into conflict with each other.

The two families of interests are distinct from the physical interests that have appeared in objective lists such as the capabilities approach to date. One might, I guess, claim that we could include these concerns in a broader understanding of Nussbaum's 'control over one's environment' or possibly even within 'senses, imagination and thought'. Where these factors are placed in any potential list is of less importance than the fact that they are included in that list. The fact that they are both families of interests, that they have a conceptual difference to the existing categories, and that they are growing in importance suggests that they are best represented as their own category rather than shoehorned into another.

Three dates were chosen earlier in this chapter to demonstrate the increasing importance of informational aspects of our lives. If a fourth date was chosen, this time in the future, it seems likely that that importance will have increased even more. As one of the project's axioms states, we are only partially through the digital revolution. Moreover, both private data control and public data access are families or clusters of

concerns and as such it seems correct to position them at a similar level to, for instance, ‘affiliation’, ‘practical reason’ or ‘emotions’ which act as holders for a range of issues. In this respect, they might be seen as headings in the list of interests rather than specific interests.

That is not to say that there is no interrelation or interdependence between these cyber interests and physical interests. The digital interests may interact with other interests in a complex fashion, and those interactions should be regarded as important and significant. Rather, as Nussbaum makes clear in the earlier quoted section, these interests (or capabilities) are ‘separate components’, and a lack in either private data control or public data access cannot be necessarily overcome by excess of other interests.

Given that a pluralistic interpretation of human interests is required, this section has demonstrated that cyber interests, which represent our informational interests, are required to reflect the growing importance of informational aspects of our lives. Prime among these informationally grounded interests are private data control and public data access.

3.5 Cyber Methods

In the conversation that surrounds cyber methods, talk about military uses of cyber ‘weapons’ often drowns out all other discussion. As was discussed in the first chapter, military cyberwar is more discussed than societal conflicts of the type that Arquilla and Ronfeldt (1993) called ’netwar’. However, as discussed in Chapter One, direct military use is only one form of cyber method. The digital revolution is a major societal change that is altering both the ways in which we interact with each other and the things that we find valuable. In the context of large-scale conflict that results in both a change of methods and a change in motivation.

One of the substantive differences between cyber actions and conventional actions is the ability of cyber actions to target informational life without necessarily targeting biological life. Although humans have always been informational creatures, methods that achieve this distinct targeting of the informational body have either not existed or have been so small scale - such as physically altering a written record - that they have not been significant.

The thoughts of the previous paragraph pick out two aspects of the digital revolution. The first is the fact that although humans have always been informational, there is an increasing distinction between their informational bodies and their biological bodies. This allows cyber methods to act on the informational body without acting on the biological body. The second is that the informational body is increasingly distributed, persistent, and accessible. This makes possible methods that target that information and results in vulnerability. Accessibility allows vulnerability and this vulnerability allows the distinct targeting of the informational body. The increasing separation of

the informational body is an artefact of the digital revolution, as is the vulnerability that it creates.

3.5.1 Separation of the Informational and Biological

The digital revolution has altered the potential separation between the informational and biological bodies. In order to demonstrate the possibility and impact of this potential, three fictional historical individuals are presented. They all live in Redmile, which is a rural village in Nottinghamshire - roughly at the geographic centre of modern England. The dates or the examples are chosen with a certain degree of care. The first, 1085, is the date of the start of the 'Great Survey' that resulted in the Doomsday Book which amounts to the first survey of its kind. In the context of large-scale conflict, it is worth noting that it was initiated by William the Conqueror who was largely concerned with discovering 'what he had' for reasons involved in the levying of taxes. The second date is 1873, a date which signified the next real attempt to perform a survey or census in the UK, this time driven by the first liberal government's concern with the distribution of property ownership. The final date is 2020, only chosen because it is the date when these words were written.

In 1085, William tasked his men to find "How many hundreds of hides were in the shire, what land the king himself had, and what stock upon the land; or what does he ought to have by the year from the shire." (Killings 1996) The resulting work - the Doomsday Book -was the first real census or national survey in the UK. However, it was severely limited and even the population of England is uncertain,

> The total population of England in 1086 cannot be calculated accurately from Domesday for several reasons: only the heads of households are listed; major cities like London and Winchester were omitted completely; there are no records of nuns, monks, or people in castles. (The Domesday Book Online 2019)

In terms of their biological lives, the occupants of Redmile in 1085 are not significantly different from the occupants in the other dates. By which is meant that they interact with the physical world in much the same way. They require food and drink in the same manner, they walk, talk and function in much the same manner. Genetically, they are largely identical to the residents of later dates. At birth, there are not major biological or physiological differences between the groups, though the prevalence of disease, malnutrition and the like will have biological effects in later life. The actual lives of the residents of 1085 are different from 2020 in many ways, but due to the speed of evolution their biological interactions with the physical world are largely similar to those of the residents of 1873 and 2020. On the other hand, their relationship with information is quite different. This is not to say that they are not informational creatures. Their innate cognitive potential is not greatly different from those of the later residents of Redmile. They do not have access to written material, or for that matter a formal education, and they do not have access to digital media and storage. The most sophisticated way that the majority of information is exchanged

is by verbal communication and this limits the scope of the information to which an individual has access. Moreover, they do not create the informational trail that subsequent residents of Redmile create. The very fact that the Great Survey was necessary demonstrates the limited accessible and persistent information that was created by individuals in 1085. For the resident in 1085, that information exists but it is not persistent. It soon is lost, and it is only accessible to very few people at the time and no people in later ages. The lack of accessible and persistent information associated with individuals meant that it was virtually impossible to wage any type of informational war or attack on them. William's worry was apparently that elements of the population would be immune from taxation simply because of a lack of information. However, beyond that is the fact that a population such as this is largely immune to informational war in - even the most basic forms such as propaganda campaigns. There is no mechanism with which to successfully distribute information, the information available to the individuals cannot be easily influenced, there is limited information over which they might claim ownership, and only a minor amount of information affects their daily life.

By 1873, things were significantly different. The introduction of the printing press and printed material had changed the environment. Individuals are surrounded by much more accessible and persistent information. Their births are recorded, as are their deaths. The legal system is developed and is based on written documents. Property ownership is documented. Contracts between individuals and organisations exist. The idea that there is a collection of information that defines or delineates an individual's life starts to take on importance. The quantity of information available to individuals, in the form of newspapers, bulletins, broadsheets, and books is greatly increased and the postal service allows people to communicate over greater distances.

The nature, type and quantity of information vary according to demographics and other factors, but the information body is starting to develop more fully than it was in 1085. Information exists about citizens and those citizens have information which, in modern terms, they might regard as their own. A typical individual may have recorded details of birth and marriage. They may have written records regarding the ownership or lease of their home - whether they know it or not. They may have contracts concerning other aspects of their lives. They may have access to information in the form of books, broadsheets, and other publications. The amount of persistent and accessible information associated with an individual has expanded greatly.

However, because of the way that information is stored, interference with it is challenging. The written nature of information means that it is persistent and accessible to a certain extent but is pragmatically immutable. There may be cases in which information was nefariously altered but widespread interference with data would be difficult.

In 2020, the situation is considerably different once again. It is clear that there is an extensive collection of information that defines our lives. We can be identified by this collection of information, and it affects us in all aspects of our lives. It defines our

wealth, our citizenship, our legal status, our affiliations, and our friendships. We have passports, birth certificates, marriage certificates and contracts, mortgages, and bank accounts. While much of that information is issued in 'paper' forms, the authoritative version of the records is increasing in digital form. That authoritative record is persistent and accessible but also mutable in a way in which a written record is not. This allows the potential for widespread informational interference. Moreover, we live our lives surrounded by public information and that information informs and affects us in significant ways. Again, increasingly that information is in digital forms, and this means that interference can be widespread and effective. For the residents of 2020, informational interference can happen both in the areas of private data control and public data access.

The emphasis in these examples has been the prevalence of accessible and persistent information in the citizens' lives. Without persistent information - in other words, without stored information owned by, created by or about an individual - the long-term influence of that information is impossible. Without accessible information, methods that interfere with that information become much less plausible.

The significant change in the 21st century is not that we have become informational creatures. It is the environment in which we operate that has changed and our relationship to that environment. The development of society and the prevalence of the written word allowed the importance of persistent and accessible information to grow. The digital revolution expanded that growth exponentially but also demonstrated that the informational was discrete from the biological. In 1085, the amount of an individual's information that existed independently of the biological body was extremely limited. By 1873, there was an amount of information that existed independently. By 2020, the vast majority of information regarding an individual exists persistently and in accessible forms and is independent of the biological body. This process, in which much of the information that defines an individual is increasingly independent of the biological body, is only likely to continue. It is this fact that enables informational conflict, that makes societal cyber conflict significant. Cyber conflict has methods of targeting individuals that did not exist previously because those things which define our lives are increasingly discrete from our biological bodies.

3.5.2 Increasing Vulnerability

The informational body is increasingly important in our lives and is at least partly constituted by persistent and accessible information. This is information with which the individual has a deep and intimate relationship and the fact that it is persistent and accessible allows it to become vulnerable. This vulnerability is best demonstrated in negative cases.

If information is not persistent then it cannot be vulnerable in this sense. Physical tracking data of the type stored on Google servers provides a clear example of this. The residents of 1085 went about their lives in Redmile creating informational detail about their location. A fellow villager might ask a friend where a particular person was.

Maybe the reply would be, 'in that field over there', but that information was neither accessible remotely nor did it persist beyond the immediate timescale. Contrast that with the case today, where my location is stored on Google's servers. This is without considering the other ways in which my specific or general location is recorded - bank account addresses, credit card use, surveillance cameras and suchlike. That information is persistent in that it will exist, for example, a year after it is created. That is largely untrue of the location information of the resident of 1085.

Additionally, if information is not accessible then it is not vulnerable in this sense. In this case, the more useful comparison is with the resident of 1873. Information about their life, which was becoming increasingly important, was predominantly recorded in written form. So, for instance, the births in Redmile in 1873 were registered locally but that information was subsequently passed to the Registrar General at the General Register Office (GRO) in London. In order to interfere with the authoritative record of a birth, one would have to access the record at the GRO and physically change it. That process is obviously not without its difficulties. In contrast, when a birth is registered in 2020, although a physical birth certificate is issued, the authoritative record is stored digitally. The process of interfering with that authoritative record is likely mediated through cyber means and is facilitated by the fact that the information has a much greater degree of accessibility. Of course, one is still going to have to circumvent security but the process of altering, for instance, 10,000 records, is plausible where it simply was not in 1873.

3.5.3 Wrapping up

Previously, cyber bullying was used as an example of how we understood the meaning of words that used the 'cyber' modifier (§1.3.1.). From there, we progressed to argue that that the word cyber references the relationships between humans and (usually digital) information. A response to this is to ask what informational bullying would look like outside the digital age. This section has provided the tools to answer that question. When the majority of information was stored in printed form, it might have been possible for me to bully a colleague by altering all references of their name to something uncharitable or offensive. In the film *Four Weddings and a Funeral*, the character Henrietta is always referred to as 'Duckface'. This is a form of informational bullying. If I alter all references to Henrietta in the information that relates to her to 'duckface' then this amounts to informational bullying. Practically, this is very hard to do in the pre-digital environment, but it is not conceptually impossible. What has changed is not that information is important to us, but that the quantity of persistent and accessible information has increased exponentially.

Because information is now increasingly persistent, it is increasingly vulnerable. Because information is increasingly accessible, it is increasingly vulnerable. Persistence and accessibility allow vulnerability.

Any method that aims to leverage this vulnerability needs to target information that plays a significant role in individuals' lives. Another way of saying this would be to

describe the target as information with which we have a meaningful and significant relationship. It is that relationship that constitutes the heart of ‘cyber’.

3.6 Two Challenges

There are two major challenges to the conception of the importance of cyber capabilities presented above that justify a more complete discussion. The first is that these cyber capabilities are dependent on physical factors and therefore are secondary to those concerns. The second is that this is a ‘first world problem’, by which is meant that much of the world’s population do not have the same attitude towards these capabilities.

3.6.1. Biological Life is Primary

The claim that has been made above is that cyber capabilities should be considered alongside physical capabilities. The sceptic will say something like, ‘if one has one’s head chopped off, then all this talk of the deep import of information is really quite irrelevant.’ It is a strong objection and deserves a proper answer. The challenge can appear in a number of forms but the thesis that underpins them all is that without biological life all these other potential aspects are evidently unimportant.

This is a challenge that is not limited to cyber capabilities. It is a challenge for the capabilities approach overall. The sceptic could just as easily have asked what value is a family of capabilities centred on emotions if one is dead? However, the inclusion of cyber capabilities does raise particular concerns and it is worth taking the sceptic’s challenge seriously. If biological life has some primacy, then should we concern ourselves with other factors, particularly factors as seemingly remote as cyber capabilities?

Although the cyber domain is only of recent importance, there is some historical precedence for this discussion. For many religious believers, the biological and spiritual have coexisted in a state of mutual reliance and importance. Both had deep import and significance in an individual's life and in their humanity. Across religions the manner in which the spiritual was related or linked to the biological varied somewhat, but a living human had both a physical body and a soul. Absence of either of these elements resulted in something that was not fully human. This tangent is introduced only to demonstrate that the idea that humanity is not only a function of biology - an idea to which we will return in a few paragraphs - is not without precedent. So, one answer to the sceptic is that, if their argument were correct, then all prior talk of the deep import of spirituality is also irrelevant. Of course, the sceptic may be irreligious and be more than happy with the bullet-biting involved in that concession.

One of the reasons that this challenge is compelling at first glance is a certain ambiguity about what is meant by life. Life can have a purely biological meaning. For example, the question ‘is the cat alive or dead?’ traditionally has only two separate

and distinct answers. A ‘human life’ can have this narrow meaning. However, the phrase ‘a human life’ can have a wider meaning that incorporates a multitude of factors that go to make up a human life. The ambiguity that I referred to above is generated by the fact that asking if an individual is alive or dead is not the same question as asking if the individual has a meaningful human life. A biological life is necessary for a meaningful human life. In the midst of the treacherous ground that surrounds the issues of what constitutes a meaningful human life, this feels like more solid ground. If one conflates the feeling of security of that solid ground with the ambiguity of the word ‘life’ it can appear that the biological is not only necessary but also sufficient.

It is abundantly clear that the biological is not sufficient. There are diverse ways in which that can be demonstrated. For the religious, a biological life without a spiritual component or soul is not a human life. A more physiological approach is to argue that when mental processes cease biological life remains a possibility with the assistance of medical apparatus. This ‘brain death’ is poorly understood in general (Jones et al. 2018) and is somewhat ambiguous and variously defined. It may be interpreted as the absence of high-level executive function - where autonomous processes are still active - or the absence of all brain activity. However, in both cases, the biological life of the body can be perpetuated with medical aid. The fact that there is active ethical discussion about the exact level of cognitive degeneration at which one should accept that life is no longer viable demonstrates that without some degree of interaction with information most of us accept that a life is not a meaningful human life. We may argue about the level of degeneration, but the concept that interaction with information is essential to a meaningful life is widely accepted and a life in medical terms is defined as much by interaction with information as it is by biological processes.

Philosophers, notably David Chalmers, are fond of examples involving zombies, or more accurately p-zombies (philosophical zombies). Robert Kirk explains the philosophical zombie,

> Zombies in philosophy are imaginary creatures designed to illuminate problems about consciousness and its relation to the physical world. Unlike the ones in films or witchcraft, they are exactly like us in all physical respects but without conscious experiences: by definition there is ‘nothing it is like’ to be a zombie. (Kirk 2021)

The p-zombie is used primarily in discussions surrounding the nature of consciousness. However, what underpins the concept is that they are biologically alive, but they interact with information in a different way to humans in that they do not have conscious experiences. What is often unsaid, presumably because it goes without saying, is that the p-zombie is not human. Or at least, not fully human. Even though they are biologically alive their existence is not a human life.

The example of the p-zombie leads to a further comment from the sceptic. ‘All you are saying,’ they might complain, ‘is that the cognitive is essential for a meaningful human life.’ In terms of the argument, this might be a tempting avenue to follow for

the sceptic because they could then claim that the cognitive reduces to the biological and the challenge might stand. Ultimately, it is fruitless for the sceptic because the claim that humans are informational creatures is significantly different from the claim that cognitive processes are essential to human life.

The cognitive processes of a living human are a primary way in which interaction with information occurs. Previously, two informational traits were outlined that described the ways in which humans interact with information. The first manner was the direct interaction with, or consumption of, information. The second was the processing of information or the creation of new information. It was stressed that this second manner was not necessarily directly related to cognitive processes. Our lives create information in ways that are not directly related to conscious thought. Think, for example, about the data that Google stores about your locations. As was demonstrated by lawsuits regarding this form of tracking (BBC News 2020), individuals may legitimately feel that this information is 'theirs'. One might argue that the information was, in fact, a result of cognitive processes - because we had consciously or unconsciously decided to go to those places - but this seems like an overly reductive approach in which all aspects of life are homogenised as ultimately reliant on cognitive processes. The challenge for an advocate of that route is that if all aspects of life are reliant on cognitive processes, then it is hard to dismiss the value of information in our lives.

Our modern lives are deeply related to information. Our qualifications, our bank accounts, our properties and possessions, our friendships and social interactions and most other elements of our life are all based on information and each time we interact with this ocean of information, we leave a trail of further information. It is this trail that largely defines our life, our history and our identity.

The challenge that, without biological life, this talk of informational life is moot fails because biological life is necessary for a meaningful human life, but it is not sufficient. Without an informational body, our lives are far from meaningful human lives. We are closer to p-zombies. Interpreting a meaningful life in purely biological terms is reductive in a way that over-simplifies, and in so doing neglects essential elements that make that life meaningful. This understanding drives us towards a pluralism that is complex, but better reflects the reality of human lives.

The claim that is being made is a social claim that there is more to a flourishing human life than a biological life. This is not primarily a metaphysical claim or a claim that is dependent on a particular theory of mind. Rather, it is a social claim that has influence as a political claim. It entails a commitment to the belief that the correct arena of social and political theory is the development of flourishing human lives, which has been presented as one of the axioms of the project.

3.6.2. It is a First-World Problem

The claim of the previous sections was that private data control and public data access are important factors in the assessment of a flourishing human life. There is an obvious challenge to this claim that is based on the non-universal nature of those capabilities or interests. 'What,' asks the critic, 'does a member of the less advantaged populations of the world care about these capabilities?' The implication is that these are 'first-world problems' that are only of concern to the privileged of the world. The challenge is that, for much of the world's population, the effects of the digital revolution are not relevant and the claim that informational interests are critical to their lives is removed from the reality of their lives.

There are three responses to this challenge.

The first is to note that other factors that we might think of as essential to a thriving life, for instance access to education or modern health care, are also not uniformly distributed on a global scale. This lack of equal dispersion does not lead us to doubt the value of these factors.

The second is that the factors can be important to their lives whether they are aware of that fact or not. Decisions that are made by local authorities, national governments, the UN, international courts, and other organisations directly affect the lives of individuals. These decisions are at least partially guided by information about individuals and populations. It is true that less developed populations are partially alienated from this information but that does not imply that the information is not important in their lives.

In addition, the value of a factor such as public data access is not determined by whether that value is realised or not. In the language of the capabilities approach, it is capabilities that are values, rather than functionings. Saying that information access and control are unimportant to groups that do not possess or value them would be equivalent to claiming that education is not of value to groups that do not have or value it. One does not need to take a universalist position in which values, interests or capabilities are shared by all of humanity, to believe that. One can believe that education provides a family of concerns that has importance to a flourishing life, while believing that the family of concerns is likely to be interpreted differently in different cultures and locations. This perhaps is closest to Sen's interpretation of capabilities.

The third response is that we are in the midst of a digital revolution rather than at the end of it. If we take the industrial revolution as a paradigm of how this type of major societal revolution may affect individuals around the world, it is clear that the effects have propagated around the world over a period of time. We can realistically expect an increasing number of the world's population to be directly affected, and concerned with, these issues.

If we believe that the majority of the world will in the short- or middle-term future regard these factors as critical, the challenge fails. Only if one believes that the divide

between those who do and do not regard information as critical to their lives will persist indefinitely does the challenge stand. Such a case, in which there is a two-tier society, is not entirely implausible. However, in this case, the awareness of the 'digital capabilities' and their absence from a section of the world's population provides a method of demonstrating the oppression that is implicit in a two-tier system. Either way, the incorporation of private data control and public data access into moral and pragmatic assessments is essential.

3.7 Conclusion

I have argued that cyber capabilities must be regarded as constituent of a flourishing human life. This argument is based on the idea of cyber interests and has used the language of the capabilities approach but might equally have been conceptualised using alternative frameworks. Degradation of any capability can result in the denial of a flourishing human life. Therefore, interference with our digital capabilities can result in the denial of a flourishing human life. Cyber methods, those that target the relationship of humans to information technology, are the way in which this fact can be leveraged. Cyber methods can act on our biological bodies and other elements of the physical world due to their transversality. They can also act on our informational bodies and in so doing are capable of denial of flourishing.

Given the fact that we are in the middle of the digital revolution rather than at its end, the exact nature of the prospective digital capabilities is somewhat anticipatory. However, two families of digital capabilities are emergent. These are ***private data control*** and ***public data access*** and constitute families of capabilities in much the same manner as 'affiliation' or 'practical reason' form families of capabilities in conventional capabilities approaches. What the language of the capabilities approach allows is the inclusion of these families of interests in the group of those things that are necessary for a meaningful human life.

Once again, it is worth emphasising that this conclusion, that ***private data control*** and ***public data access*** are necessary components of a truly meaningful human life, is independent of the capabilities approach. It might be described equally using either other objective list approaches or other theories of welfare. In these cases, the growing importance of private data control and public data access would still need to be reflected. The value of the capabilities approach is that it provides both appropriate and simple language and a pluralistic bedrock that aids the discussion.

Cyber methods are particularly relevant in this context for two reasons. The first was that our informational body is increasingly discrete from our physical bodies which allows interference of the informational body without interference with the biological body. The second was that information is increasingly accessible. These two reasons combine to determine that our informational bodies are increasingly vulnerable to interference. Any interference would directly influence the relationship between humans and information. It is this relationship that defines 'cyber' and, therefore, cyber methods are the prime manner in which the informational body may be targeted.

The aim of this chapter has been to construct a framework of analysis that is appropriate to the post-digital-revolution world. The use of a pluralistic interpretation of harms and interests, including cyber interests, has been shown to be appropriate. In the introduction, I stated that we would have to revisit the assumptions on which this conclusion is based. The general assumptions are included in the axioms of the project. *We value flourishing human lives*, *harm as a degradation of flourishing*, *we fight about those things we value,* and *we are in the midst of a digital revolution*. All those axioms have had some bearing on this chapter but perhaps the most significant is the fact that flourishing human lives are the correct quanta of evaluation, and that harm may be represented as a degradation of flourishing.

This project is inextricably linked to those axioms and in the case of the value of flourishing human lives as the correct quanta, I have little more to say. If a sceptic denies that that is the case, then I fear that we will simply part ways. That harm is represented as a degradation of flourishing is perhaps a less engraved axiom. There are other valid ways of representing harm. I am open to that possibility but do not believe that any alternate representation would diminish the fact that a degradation of interests does constitute harm. There may be other ways of expressing this, but the fundamental truth remains. It does not seem that the conclusions of this chapter are based on any framework or ethical system other than the original axioms.

We are left with a framework that we can use to evaluate harms and interests that is based on very few assumptions. This framework will be used in the following chapters to demonstrate that the current ethical theories of war and large-scale conflict are simply unable to cope with the challenges presented by the post-digital-revolution world. Where does this leave us with regard to the scope of the existing claims, and the claims that are going to be made based on this framework?

With regard to cyber issues, there is a perennial question as to whether one is overstating the matter. In this analysis, the claim is that in the post-digital-revolution world cyber interests take a position that is equivalent and equal in significance to our other interests. They cannot be written off as simply ones-and-zeros because we are intrinsically informational in our nature and the digital revolution has simply allowed that aspect of our existence to increase. A complete description of a flourishing human life now, and in the future, will need to accommodate informational and cyber interests. Is the description of the significance of cyber interests presented here overblown? Not at all.

Conversely, it might be claimed that this analysis does nothing other than scratch the surface of the societal changes that are likely to occur. In limiting the analysis to large-scale conflict, the most important elements of the potential discussion are denied. Political philosophy is largely concerned with the relationship between the individuals who are governed and the system that governs them. Some readers may feel that the conclusions drawn here regarding large-scale conflict are only a side-effect of larger societal changes that will inevitably influence the relationship that individuals have with systems of government. It seems hard to deny that possibility. The argument goes something like this. What individuals hope to gain from systems of government

is an environment in which they may flourish. If the concept of flourishing is changing, then the systems of government must change. The impact of this on the current project is that it may be that the changes of the digital revolution are so great that the subsequent analysis, based as it is on current ethics of conflict, is likely to be quickly obsolete. In other words, the concern is not that the scope is overblown but that it is under-blown.

The under-blown argument is at least potentially persuasive, and with that in mind the final chapter of this project, ‘Subversion and Just Cause’ is more speculative than the main body of the project which applies the new framework to the main elements of JWT. Major, and necessary, constituents of moderate theories of war are shown to be unable to accommodate the conclusions of the new framework. Chapter 4, *Sabotage and Discrimination*, undermines the principle of discrimination that is critical to our intuitive acceptance of JWT. The following chapter, Espionage and Proportionality, takes the same approach and shows that proportionality is not tenable in the post-digital-revolution world. Without discrimination and proportionality is hard to see how any moderate theory of war can remain appropriate.

People sometimes ask whether ‘cyber’ is really that different. The answer is yes, and it is demonstrated here. One may try mistakenly to dismiss ‘cyber weapons’ as simply another weapon if one thinks only in terms of methods. However, if one starts one’s analysis from the concept that the digital revolution is changing those things that we value, the conclusions are that key aspects of moderate theories of war become untenable.

Interlude

Part One of this project (Chapters 1-3) created a framework that stressed the developing importance of cyber interests in the context of a pluralistic interpretation of harm. Two cyber interests were described that have significance in the assessment of flourishing human lives: private data control and public data access. A 'cyber interest framework' is created that is used in the following chapters. A summary of the argument that got us this far is presented in the next three paragraphs.

In the first chapter, the idea was developed that moderate theories of war such as JWT attempt to balance the harms and the benefits of large-scale conflict. This forms one side of the landscape of this investigation. The other side is the nature of cyber conflict. ‘Cyber’ was defined as concerning the relationship between humans and information technology. From the idea that it concerns a relationship, it followed that cyber conflict encompasses both a change in methods (cyber methods) and a change in the goods that are at the root of conflict (cyber interests).

The predominant interpretation in theories of war has been based on a perception of war as defined by killing. In the second chapter, it was argued that this perception was inappropriate to modern large-scale conflict. Not only did its use result in poor and sometimes incorrect analysis, but it could also be leveraged by unscrupulous actors to justify the imposition of harm. Use of a single metric such as killings to evaluate a complex phenomenon such as harm is inappropriate.

In the third chapter, a pluralistic appreciation of harm was presented. In the context of large-scale cyber conflict, it was suggested that public data access and private data control formed two families of interests that are important and likely to become increasingly important in people’s lives. In terms of cyber conflict, consideration of all human interests including these cyber interests is necessary if one is going to evaluate the moral significance of actions. Without inclusion of the entire plurality of interests, any assessment becomes inaccurate and therefore morally questionable. Although we are at a point in history where cyber conflict is being developed, two families of cyber interest were suggested: control of private data and public data access.

The second part of this project (Chapters 4-6) takes the pluralistic interpretation of harm and applies it to essential elements of moderate theories of war such as JWT in the context of cyber conflict. The main concept that underpins the structure of Part Two is a novel concept of direct/system influence. This concept distinguishes actions on the basis of their influence. Direct influence is the type of influence that acts directly on an asset. So, for instance, blowing up a dam is an example of direct influence on a physical asset. System influence is the type of influence that affects the asset by

acting on its systems of production and control. So, for instance, disabling a dam by interfering with its control systems is an example of system influence on a physical asset.

There is an area of potential confusion that needs to be resolved. Control and production systems may exist in either the physical or informational domain. For instance, factories may be part of the production systems of machinery such as tanks and exist in the physical domain. There are also control systems associated with a modern factory and those exist in the informational domain. The categorisation that is used here is based on the effect of the system. If the effect is primarily physical - as in the case of a factory's control systems - then that is categorised below as physical. If the output is informational - as is the case in the Domain Name System (DNS) that provides routes for Internet messaging - then that is categorised as informational. The primary ***effect*** is the method of categorisation.

The difference between direct influence and control and production system influence emerges during Part Two and is clarified as we go along, but it is helpful to have it in mind from the start. It forms one axis of a table that is shown below. The other axis is a distinction of whether the main effects of the action are categorised as occurring in the physical domain or in the informational domain.

INFLUENCE TABLE	Direct Influence	Systems Influence
Physical Effects		Sabotage (Discrimination)
Informational Effects	Espionage (Proportionality)	Subversion (Just Cause)

The table is valuable in navigating the next chapters but also interesting straight out of the box. There is a blank square in the direct influence/physical domain box. This might be expected as cyber methods are not suited to direct influence on physical assets. The structure of a castle is at negligible risk from cyber-attack, even if the systems that manage the operation of the castle may be at risk.

Thomas Rid famously claimed that the majority of cyber conflict could be categorised as sabotage, espionage, and subversion. Both the empirical evidence of the following years and further studies (e.g., Valeriano et Al. 2018) have done nothing to dispel the accuracy of these categorisations. However, scant attention has been paid to the conceptual underpinnings of those categories. The table above allows those categorisations to be more fully understood in terms of cyber actions. Cyber sabotage is an action that has physical effects brought on by influence on systems controlling the asset. Cyber espionage is an action that has direct influence by stealing or modifying an informational asset. Cyber subversion is an action that has influence on informational systems of control and production. One can be confident that the table above fully covers the conceptual space and so can be confident that discussion of sabotage, espionage and subversion does so too.

Happily, the tripartite structure of the categories of cyber action (sabotage, espionage, and subversion) is matched by the elements of moderate war theory that are most valuable to discuss in the context of cyber conflict, the principle of discrimination, the principle of proportionality and the principle of just cause. Each of the categories is paired with one principle as is explained below. It is worth stressing that the pairings are an artefact of the tripartite structure and should not be taken as overly significant. For example, even though discrimination is paired with sabotage, this does not mean that discrimination is not appropriate in the other categories. Nor does it mean that the other principles do not apply to sabotage. The pairings are simply artefacts that may make the scale of the subject more easily digested.

In Chapter 4, the consideration of sabotage starts with a seemingly simple point; that cyber interference with the physical world is often best achieved by interfering with systems of production. Cyber methods can have little effect on steel itself, but they can have significant effects on systems of control relating to steel production. The division between the asset itself (direct influence) and the systems involved in its production and control (system influence) is the division that emerges as significant throughout Part Two. Regarding effects in the physical domain, cyber methods are ineffective on direct influence and highly effective on system influence.

The fact that cyber sabotage has effects on systems of production in the physical domain results in it forming a compelling pair with the principle that demands that those effects should be inflicted only on those who are liable to them - the principle of discrimination. Sabotage provides appropriate examples for the discussion of discrimination. Cyber sabotage is the only category in which the effects are largely in the physical domain, and this means that this chapter is the one in which the cyber interest framework of Part One is of least value. The true value of that framework appears in the next chapter.

In Chapter 5, the consideration of espionage strengthens and clarifies the concepts of direct influence and systems influence. Data is an informational asset. Cyber

espionage is a cyber method that exerts direct influence on that asset. It does not involve system influence; in that, it does not act on systems of production of that asset. Because espionage highlights aspects of actions against assets it pairs nicely with discussion of proportionality. The problem that forms the key element of this Chapter is how one may balance the worth of informational assets against the worth of physical assets. The cyber interest framework demands that informational assets must be accounted for when making ethical assessments of actions.

Chapter 6 pairs Just Cause with subversion. Subversion is conceptualised as an attack on systems of control and production in the informational domain. In particular, subversion is an attack on high-level organisational structures and systems that are informational in nature. As such, the state is considered an informational structure. The most widely accepted instance of just cause is national defence and subversion is interpreted as an attack on the conceptual structures of a particular nation. The use of a control systems framework coupled with the understanding of Part One leads to a theory that accommodates the right of national defence which in other theories is more difficult to justify. However, there are also strong challenges to the very idea of the nation that emerge from this control system framework in the context of the cyber domain. These challenges hint at ways in which a post-digital-revolution society might be arranged. It is outside the scope of this project to resolve all of these challenges but the potential of the analysis to this point hints at further use.

The primary aim of this project is to demonstrate that moderate theories cannot be applied in a principled manner in the cyber context. With regard to this aim, much of the heavy lifting is done in Chapters 4 and 5. Without some implementation of proportionality and discrimination (or at the very least a reflection of the pre-theoretic thoughts that motivate those principles), a moderate theory cannot stand. Chapter 6 is different for two reasons. The first is that the previous two chapters have used the cyber interests framework in a negative manner and Chapter 6 aims to demonstrate that the framework can be used in a more positive fashion - to elucidate some of the questions around the thorny issue of national defence. New and potentially positive understandings of civil relationships and society overall may be created by the application of this framework. The second is that Chapter 6 does not directly focus on the primary aim of the project. Rather, it focuses on some of the challenges and questions that arise from the conclusion that moderate theories cannot be applied in cyber conflict. As was discussed in the introduction, if one reaches this conclusion one is faced with a choice between realism and pacifism. Chapter 6 might be regarded as an initial investigation of where the cyber interests framework might take one in that discussion. In particular, it aims to provide some evidence of the challenges that will lie in store for the realist route.

4. Sabotage and Discrimination

Sabotage did not involve loss of life, and it offered the best hope for future race relations. Bitterness would be kept to a minimum and, if the policy bore fruit, democratic government could become a reality.

Nelson Mandela (1964)

4.1. Introduction

The harms involved in large-scale conflict are morally significant but despite this fact, moderate theories of conflict allow that there may be justifications for those harms. One way of phrasing this might be that harm may be imposed on people if the overall benefits are sufficient. A question that emerges from this phrasing is, 'which people?' That is the question with which the principle of discrimination is concerned, and it is a principle that aims to distinguish legitimate targets on which harm may be imposed from those who should be immune to that imposition of harm.

There is no single definition of discrimination. The principle was easier to understand when war consisted of highly organised and uniformed soldiers on a defined battlefield. In those conditions, it was more straightforward to distinguish justifiable targets. Modern warfare has become more complex than that idealised image of war and defining who should be designated as legitimate targets has become correspondingly more difficult. This is the conceptual question of discrimination (e.g., Primoratz 2007 & Lazar 2015) that aims to define those factors that contribute to making an individual liable to harm.

However, there are also epistemic problems involved in assessing situations with regard to the principle of discrimination.

Firstly, even if we have a conceptual idea of which individuals may be targeted, are we in an epistemic position to separate real populations into those who may be targeted and those who may not? For instance, even if we are happy to accept a civilian/military divide, can we be sure of who is a civilian? I refer to this as the epistemic problem of recognition.

Secondly, even if we have a conceptual idea of which type of individuals may be targeted, can we be certain that these are the people who our actions will affect? Even if we can satisfy the recognition requirements, are we assured that our actions will target those individuals and not others? I refer to this as the epistemic problem of accuracy.

Both types of epistemic problem can undermine any conceptual formulation of discrimination that aims to be action-guiding rather than purely conceptual. Perhaps we must be tolerant of some epistemic worries. Total certainty is an exceedingly high bar indeed and, particularly in the context of large-scale conflict, may be obscured by 'the fog of war.' However, the epistemic problems centred around cyber methods are central rather than marginal. They stem, not from a generic 'fog of war', but from the actual characteristics of the methods themselves. As such, the application of discrimination in cyber conflict becomes at best deeply problematic, and more likely inappropriate.

As has been discussed, it is not the aim here to resolve the issues that arise from that conclusion. The aim of this chapter is to arrive at that conclusion – that the use of the principle of discrimination is inappropriate in the cyber context. There are alternative ways in which one might respond to that conclusion. One might suggest that if cyber conflict can never meet the requirements of discrimination, then cyber conflict must not be permitted. One might, on the other hand, claim that if the principle of discrimination cannot be successfully applied then it no longer provides a meaningful limitation to what is permissible, and no such limitation exists – or anything goes. Settling these responses to the conclusion is left to later chapters when there will be further evidence available. The purpose of this chapter is to start amassing that evidence by demonstrating that it is not possible to apply the principle of discrimination to cyber conflict in a principled manner.

Cyber sabotage is used as a foil for this analysis of the principle of discrimination. It is worth stressing once again that this does not suggest that only sabotage is relevant to discrimination. Rather, its use is an artefact that relies on the fact that the pairing of sabotage and discrimination allows certain aspects of the effects of cyber conflict on moderate theories to be highlighted.

For the project overall, sabotage is a good place to start because its effects are predominantly in the physical domain. The arguments of this chapter rely less on a commitment to the cyber interests framework developed in Part One. Cyber methods can cause problems for moderate theories of war even when we only consider physical effects. This provides a solid base for the later arguments that there are increasing problems when the consideration of cyber interests is added to the mix.

4.2. Discrimination

Moderate theories of war aim to justify the infliction of harm involved in conflict by balancing the benefits against this harm. Any theory that aims to justify the infliction of harm has some intuitive components. One of those components captures the general, pre-theoretic idea that the harm should only be inflicted on the right people. The harm should not be indiscriminate. This component is called the Principle of Distinction or the Principle of Discrimination. Neither term, distinction nor discrimination, is without problems. I will use 'discrimination' only because many of the other authors quoted do so, and to do otherwise complicates understanding. This

use clarifies that actions that do not meet the requirement of the principle of discrimination can be categorised as indiscriminate.

Discrimination traditionally appears in the list of Jus in Bello criteria - the criteria of how one must wage a war for that war and those actions to be permissible. Its absence from the Jus ad Bellum criteria - the criteria as to whether we are permitted to engage in war - is an interesting fact to notice in passing. It seems plausible that an assessment of the fact we are targeting the correct people might be a criterion of Jus ad Bellum. The traditional Jus in Bello principle of discrimination concerns itself with how the war is fought – how specific targets are assessed. There is also a wider concern as to whether we are engaging with the correct opponent. A Jus in Bello principle might be concerned with which Canadians are liable to be targeted, while a Jus ad Bellum principle might concern itself with whether Canada is the correct opponent. It may be that this has traditionally been rolled into the other criteria of Jus ad Bellum or simply taken for granted. In modern conflict, the second concern is becoming increasingly significant as it is not always clear who the 'real enemy' is, and this uncertainty is demonstrated most clearly in cyber conflict. Regardless of these passing thoughts, discrimination is an analysis of whether the right people are affected by our actions.

4.2.1. The Conceptual Problem of Discrimination

The demarcation of the ‘right people’ is the cornerstone of discrimination. The prevalent way of conceptualising this issue in modern theory is in terms of liability. In a perfectly discriminate scenario, only those who were liable to harm would be harmed. The moral implications of this are that, broadly speaking, indiscriminate actions - those in which harm is applied to those who are not liable to harm - are to be avoided.

Liability is a complex issue that can be formulated in numerous ways.[19] In addition, the ethicists' understanding of liability may not match the lawyers' understanding of liability, or for that matter the layperson's use of the word. Jeff McMahan developed his idea of liability over several papers and the 2009 book, *Killing in War.* His formulation of liability has become, if not uncontested, at least a sort of baseline for following discussions. McMahan’s interpretation is that moral liability to a particular action equates to a judgement that an agent would not be wronged if that action were performed.

---

[19] For instance, see Tadros (2012) for a critique of McMahan's standpoint. See Parry (2105) for a concise summary of the discussion. "The key debate among theorists of self-defence concerns the correct rendering of the ‘relevantly implicated’ clause for grounding liability For a representative sample, see Thomson (1991), Ferzan (2005), McMahan (2005), Quong (2012) and Tadros (2012)."

In 2005, in a paper titled *The Basis of Moral Liability to Defensive Killing*, he starts the analysis in the following way,

> In other instances of permissible killing, however, the justification appeals to more than consequences. It may appeal to the claim that the person to be killed has acted in such a way that to kill him would neither wrong him nor violate his rights, even if he has not consented to be killed or to be subjected to the risk of being killed. In these cases, I will say that the person is liable to be killed. (McMahan 2005:386)

He ends up suggesting what he calls the 'responsibility account.'

> According to what I will call the Responsibility Account, the criterion of liability to defensive killing is moral responsibility, through action that lacks objective justification, for a threat of unjust harm to others, where a harm is unjust if it is one to which the victim is not liable and to which she has not consented. For convenience, I will sometimes abbreviate this to ''responsibility for an unjust threat.'' (McMahan 2005:394)

McMahan's liability is widely understood and to a certain extent forms a modern status quo. In the arguments that follow, the conclusions are not reliant on the grounding of liability that one endorses or whether one endorses liability as a metric for discrimination. A generic formulation of discrimination based on liability, that broadly equates to a responsibility for threat, is used to facilitate the conversation, rather than endorsing any specific grounding of liability.

Liability provides a functional and current way of discussing discrimination. The conceptual problem of discrimination can be defined as defining how and why certain individuals or groups of individuals are liable to harm. The fact that it may be groups of individuals who are liable raises a question as to the granularity of liability. Individualists in this context will believe that liability is limited to the individual who is responsible for the threat. Those of a more collectivist leaning may think that collective liability is possible.

If Mars poses a legitimate threat to our safety and welfare, are all Martians liable to harm or are only the specific Martians who are responsible for the threat liable to harm? If the answer is 'all Martians,' then we owe a principled answer to the question of why an individual Martian is liable simply by being a Martian - even if they are personally not responsible for any threat and may even be working against that threat. If the answer is 'specific Martians,' then the problem becomes how we define the distinction between liable Martians and non-liable Martians.

The claim was made above that the conclusions are not reliant on the grounding of liability that one endorses. This agnosticism relates to the way in which liability is defined rather than the granularity with which it is applied. A strongly collectivist interpretation of liability weakens the concept of discrimination which may explain why many writers deal in individualist terms. For instance, Ryan Jenkins says,

> An act of violence is more discriminate as the extent to which it harms particular individuals more closely approximates their individual liability. An act of violence is maximally discriminate if the harm that it causes befalls only those liable to attack. (Jenkins 2016:92)

For this project, it is the individual approach where the epistemic problems that are characterised by cyber conflict have the greatest traction. Strong advocates of rough granularity, in which broad group responsibility is accepted, may feel that the problems that are presented for the principle of discrimination may be avoided by that route. This, however, is not unique to cyber conflict. The more one relies on group responsibility, the less relevance the principle of distinction can have. If one thinks that all Martians are responsible for the threat and therefore liable, distinction becomes irrelevant. The epistemic problems discussed below have most traction on an individual approach to distinction but will have some traction on any reasonable granularity that does not make discrimination irrelevant.

There is another complication of liability that is hinted at by the quote above. In the phrase 'more closely approximates their individual liability' there is the suggestion that liability is not a binary division, or that an individual or group may be partially liable.

Traditionally, there were two categories of people that were important in this issue of discrimination: soldiers and civilians. If, for instance, I had been involved in a Napoleonic campaign as a gunner I could aim my 8-pounders at the enemy forces because they were liable to the intended harm, but could not aim them at civilians. Or that was the idea anyway. There were two categories of liability in this understanding - liable and non-liable - and it was usually trivial to divide potential targets between the groups. Those who were soldiers were liable to harm and those who were not soldiers were not liable to harm. The demarcation was usually helped by the fact that soldiers wore uniforms designating them as soldiers (e.g., Kutz 2005). Problems for this binary division arise quickly as described below and start with the fact that, as McMahan notes, the term 'soldier' is most unspecific (McMahan 2009).

The modern usage is to talk in terms of combatants and non-combatants, although the term 'civilian' appears durable, and is still widely used (e.g., Lazar 2015). For instance, a major focus of the discussion around discrimination is 'civilian immunity'.[20] In addition, there may not be a complete correlation between civilians and non-combatants. Non-combatants may include injured combatants or combatants who have surrendered. In general, again for simplicity, I will try to talk in terms of combatant and non-combatant where necessary but will avoid any binary division where possible.

It seems uncontroversial that some people are more liable than others even if they remain liable in some respect. McMahan (2005:395) has hinted at conceptual

---

[20] For recent overviews of this literature see: French 2018, Frowe 2014, and Barry & Christie 2018

underpinnings for this when he says that “The Responsibility Account treats liability as a matter of degree.” Some people may feel that the civilian manufacturer of weapons intuitively seems less liable than front-line combat troops, although his contribution to the war effort may make him somewhat liable. Others may feel that the manufacturer is more liable that the front-line troops. Exactly how that thought is developed depends on one’s underlying concept of liability, but it is enough for the current purposes that a non-binary interpretation of liability is plausible. This entails that liability is scalar and the idea of one person being more liable than another while both are liable to some extent is coherent. No greater commitment is needed than that.

Moving forward in this chapter, the principle of distinction will be understood as demanding a correlation between liability to harm and the harm inflicted. An act is more discriminate if there is a greater correlation. The problems that are discussed for the principle of discrimination are epistemic in nature and would exist irrespective of the formulation of liability that was used – or for that matter if one used a conceptualisation of discrimination other than liability. This allows a broad understanding of liability to be used in which it is interpreted as a responsibility for an unjustified threat. That responsibility is not binary in that individuals may be less responsible than others while retaining some responsibility, and therefore liability.

4.2.2. Epistemic Problems of Discrimination

If discrimination is to form an essential element of any theory of war, there are some fundamental requirements.

- We must be able to distinguish how liable certain individuals are - this is the epistemic problem of recognition
- We must be assured that the methods we use are capable of targeting with sufficient accuracy those who we recognise as liable to the harm which is inflicted - this is the epistemic problem of accuracy.

The two requirements are independent of the conceptual problem of discrimination. In the simplest terms, in which we (perhaps incorrectly) decide to divide by soldier or not soldier, the first requirement demands that we can recognise who is a soldier and who is not. In more complex and non-binary versions of liability, it demands that we can correctly assess the liability of given individuals.

The second requirement demands that we are assured of the accuracy of our methods. Given that we know who we can target and how much harm we may inflict, are we confident that our actions will match those demands? Importantly, will there be collateral damage? Collateral damage suggests that the harm inflicted has not been directed only at those who are liable to it.

This chapter demonstrates that in the context of cyber conflict both these requirements, recognition and accuracy, cannot be met. This is not because of what

might be termed ‘normal uncertainty,’ but because the characteristics of cyber methods result in inherent inabilities to meet the requirements.

We return to the epistemic problems in Sections 4.4 and 4.5. The next section lays the groundwork by analysing the nature of cyber sabotage. It outlines the specific inherent characteristics of cyber methods that cause the problems for a principle of discrimination. It is not a complete discussion of cyber sabotage, nor are cyber espionage and cyber subversion irrelevant to discrimination. Nonetheless, cyber sabotage provides sufficient examples to demonstrate the inability of any principle of distinction to accommodate cyber conflict.

4.3. Sabotage and Cyber Sabotage

4.3.1. Sabotage as a Political Instrument

Looking at the Influence Table, sabotage is defined as actions that have systems influence where the effects are physical. In contrast, if the effect of an action were simply to destroy machinery, then that would not necessarily qualify as sabotage. Sabotage requires that the target is not only the machinery but the system of which the machinery is part. If I steal a wheel from your car in order to sell it, that is not sabotage. If I steal a wheel from your car in order to prevent you from getting to work, then it is sabotage.

INFLUENCE TABLE	Direct Influence	Systems Influence
Physical Effects		Sabotage (Discrimination)
Informational Effects	Espionage (Proportionality)	Subversion (Just Cause)

As is shown by the opening quotation from Nelson Mandela, sabotage has been used as a political instrument. In the wider statement at his trial, Mandela makes clear that they only resorted to sabotage after having exhausted the possibilities of non-violent protest. It is clear that Mandela felt that sabotage was violent. Later in his statement (Mandela 1964), he says, "On the other hand, in view of this situation I have

described, the ANC was prepared to depart from its fifty-year-old policy of non-violence to this extent that it would no longer disapprove of properly controlled sabotage."

Because the definition implicit in the Influence Table is technical it is worth looking at the dictionary definition of the word ‘sabotage’ to assure ourselves that it is not contradictory,

> The malicious damaging or destruction of an employer's property by workmen during a strike or the like; hence gen. any disabling damage deliberately inflicted, esp. that carried out clandestinely in order to disrupt the economic or military resources of an enemy. (OED Online 2022c)

One general legal definition is,

> Sabotage is the act of hampering, deliberating, subverting, or hurting the efforts of another. It is most often an issue in the context of military law, when a person attempts to thwart a war effort, or in employment law, when disgruntled employees destroy employer property. (U.S. Legal n.d.)

Sabotage is not only a political instrument. Rather, it is a technique that uses damage of equipment, infrastructure, technology, and other materials to interrupt the normal systems of the target organisation.

A well-known and paradigmatic example of sabotage in large-scale conflict is the Allied attempts to interrupt the German production of heavy water in WWII. Heavy water was considered essential for the German plans to develop a nuclear bomb. For the purposes here the campaign is divided into four parts.

The first part consisted of two stages, Operation Grouse and Operation Freshman. Operation Grouse involved infiltrating a party of Norwegian commandos into the area by parachute. Operation Freshman, the second stage, involved infiltrating British engineers by glider to meet up with the Norwegian Commandos. Operation Freshman went horrifically wrong, with the aircraft crashing. Those who did survive the crashes were captured by the Gestapo, tortured, and then killed.

In the second part, Operation Gunnerside, a reinforced Norwegian team of commandos was successful. The team entered the target power station and successfully placed explosives which resulted in extensive damage that halted the production for several months.

Production was eventually restored. The third part of the campaign involved Allied bombing raids that led the Germans to abandon the projects though some heavy water that had previously been produced still remained. A final episode in the story, part four in the arbitrary designation used here, is the sinking by Norwegian forces of the ferry being used to transport the remaining heavy water to Germany.

The ‘heavy water sabotage’ provides an archetypal example of sabotage. The aim was to interfere with the German nuclear programme. To do this the system of production of an essential element of that program was targeted. The Germans had a system that they were using to progress towards the production of a nuclear weapon and the aim of the sabotage was to interfere with that system.

When the whole campaign is regarded as a campaign of sabotage it raises some interesting questions about the nature of sabotage. The aim of the first and second parts - which involved commandos - is evidently sabotage. It is the attempted interference with the correct operation of a system of production.

The third part of the campaign - involving bombing raids - leads to a salient conclusion. Sabotage does not need to be covert. This makes sense. If my aim is to sabotage your trip to work, then I might let your tyres down. Even if I make it clear that I have done this, it still would qualify as sabotage. This is not altered by the fact that conventional military forces are involved. The bombing raids which were clearly aimed at disturbing production can be categorised as sabotage. A covert nature is typical of sabotage but is not necessary.

Sabotage is defined as actions that have systems influence where the effects are physical, and this definition provides accuracy in the categorisation of actions, whether they are military or not.

4.3.2. Cyber Sabotage

How does this definition of sabotage translate into an understanding of cyber sabotage? And why are cyber methods particularly suited for sabotage?

Thomas Rid makes a pertinent point regarding sabotage when he says,

> The higher the technical development and the dependency of a society on technology (including public administration, the security sector, and industry), the higher the potential for both violent and non-violent sabotage, especially cyber-enabled sabotage. (Rid 2013:56)

A highly technologically developed society has an elevated level of integration of production and control systems. Technologically developed societies are potentially vulnerable to system influence. That technology that is targeted might be a computer or a complex element of machinery, or it may simply be a road or other element of infrastructure because those too may form essential elements of production and control systems. Preventing the technology from working in the intended fashion is the modus operandi of sabotage.

INFLUENCE TABLE	Direct Influence	Systems Influence
Physical Effects		Sabotage (Discrimination)
Informational Effects	Espionage (Proportionality)	Subversion (Just Cause)

Rid continues,

> Sabotage is the deliberate attempt to weaken or disable an economic or military system. All sabotage is predominantly technical in nature, but it may of course use social enablers. The means used in sabotage may not always lead to physical destruction and overt violence. Sabotage may be designed merely to disable machines or production processes temporarily, and explicitly to avoid damaging anything in a violent way. (Rid 2013:56)

This is consistent with the idea that sabotage has physical effects because not all effects are destructive or violent. Rid has a strict interpretation of what violence is. It is the same definition of violence that he leverages to argue that cyber conflict should not be considered as war. It, of course, is not the only interpretation of violence as the quotes from Mandela at the start of the chapter show, but his usage is not necessarily wrong - just one of several interpretations of the word.

Regardless of this, cyber methods do appear to be particularly suited to sabotage. As Rid (2013:53) acknowledges, “malicious software and cyber-attacks are ideal instruments of sabotage.” The reason that cyber methods are ideal instruments is that elements of production and control systems increasingly exist in the informational domain and yet have effects in the physical domain. ‘Cyber’ for this project refers to the relationship between the physical and the informational. Therefore, those systems are best defined as cyber systems. The digital revolution has changed the way that those systems are implemented.

A factory makes a useful generic example. With modern control systems, it is evident that the factory is a combination of informational and physical processes. Computerised control systems may exist on a number of scales within the factory - perhaps on a localised element of machinery or on a more global scale with regard the factory. Such a factory is very much a cyber entity in that it uses the relationship between information and physical processes to create a desired result. In a parallel to the argument presented in Chapter 3, the factory before the digital revolution included control systems - these control systems may have been managed by human

skills and knowledge to a greater extent than they are now, but they still existed. The digital revolution has made those systems more vulnerable to interference by system influence.

In the context of large-scale conflict, the aim of cyber methods is to interfere with opponents' control and production systems. Sabotage is system influence that aims to have physical effects and cyber methods are particularly appropriate because of their transversality. Because the target systems increasingly exist in the informational domain or rely on the informational domain, cyber methods are particularly suited to sabotage. This is the functional reason cyber methods make good tools for sabotage.

As well as the functional reason, there are political reasons why cyber sabotage my appeal to adversaries involved in conflict because they may be perceived as low risk for three reasons. Firstly, the covert (or non-overt) nature of cyber methods may allow involvement in actions to be denied and this may avoid repercussions. Secondly, the aggressor does not risk the lives of any of their own citizens or armed forces. In the modern world, it is becoming increasingly difficult for leaders to justify these kinds of deaths. Thirdly, cyber methods may be regarded as low risk because they are regarded as 'short of war.' Although some nations have overtly claimed that they may regard cyber methods as justifying a physical response involving the use of armed force there is not yet an example or precedent for this. The empirical evidence suggests that escalation from cyber methods to more extreme conflict, either in the cyber domain or the physical domain, is not a developing pattern (Valeriano et Al. 2018). Unscrupulous actors may feel at this point in history that cyber sabotage is a method of coercion and signalling that can be used with little risk.

The focus on production and control systems emphasises another element discussed in Chapter 3. The effects of cyber sabotage are not limited to the military but are more likely to be distributed across populations. For instance, there exists a significant worry that in our modern networked societies it might be possible for major elements of our infrastructure to be sabotaged. Whole populations may be affected rather than just the military. The commonest concerns centre around the generation and transmission of electricity.

> While investigating a power outage in Ukraine from last December, they found a new type of malware that appears to have been custom-made for infecting electric utilities. The malware, called CrashOverride, is fairly significant in the world of criminal hacking because its sole purpose is to sabotage a utility's operations and trigger a power outage. (Gregg 2017)

Widespread outages are the mainstay of these worries, but much of the rest of our infrastructure is also vulnerable including transport, health, education, and government. For those of us who live close to nuclear power stations, there are particular fears in this area.

This section has argued that sabotage is system influence that has physical effects. Cyber methods are particularly suited to this kind of influence. The next section provides three real-world examples that support this conclusion.

4.3.3. Real World Examples of Cyber Sabotage

Sabotage can be political, but it can also be personal. Imagine that your application for a job was turned down. You are angry and want to make your point. You want to strike back at the injustice of it all. This is how Vitek Boden felt when he failed to be hired by the Maroochy Shire Council (e.g., Smith 2001). The Shire of Maroochy is a local government area on the Queensland coast in Australia.

Boden had previously been employed by Hunter Watertech, an Australian firm that had installed the sewage equipment for the Maroochy Shire Council. He used his insider knowledge and some stolen equipment to take control of 150 sewage pumping stations. He succeeded in spilling hundreds and thousands of litres of raw sewage into local parks, rivers, and the grounds of a Hyatt Regency hotel.

This is a clear case of cyber sabotage. A control system was interfered with to produce physical results. The cyber methods used allowed the perpetrator to exert influence - that extended to 150 individual pumping stations - that would have otherwise been challenging. There are several salient points in the Maroochy narrative. Firstly, the perpetrator has inside information, and this is widely acknowledged as being the most effective factor in terms of sabotage potential. Secondly, the effects were obviously transversal in that the systems compromised were digital, but the effects were physical. Thirdly, the effects were dispersed rather than specific. The entire community including local flora and fauna were harmed. In this respect, the issue highlights the range of use of the word ‘harm.’ Rid (2013:78-79), for example, claims that humans were not harmed, by which he means that humans were not physically injured. This is clearly a narrow definition of harm as most would agree that having thousands of litres of sewage pumped into one’s neighbourhood was harmful. Fourthly, the actions were initially covert even though subsequent attribution was simple.

The Shamoon attack (Johnson n.d. & U.S. Government 2016b) was a more highly organised instance of sabotage used as an implement of international relations. The target was Saudi Aramco, a national oil company and at the time earned the title "the biggest hack in history" (Pagliery 2016). The attack was announced on Pastebin, which is an anonymous forum, on August 15th, 2012,

> We, behalf of an anti-oppression hacker group that have been fed up of crimes and atrocities taking place in various countries around the world, especially in the neighboring countries such as Syria, Bahrain, Yemen, Lebanon, Egypt and ..., and also of dual approach of the world community to these nations, want to hit the main supporters of these disasters by this action.

> One of the main supporters of this disasters is Al-Saud corrupt regime that sponsors such oppressive measures by using Muslims oil resources. Al-Saud is a partner in committing these crimes. Its hands are infected with the blood of innocent children and people.
>
> In the first step, an action was performed against Aramco company, as the largest financial source for Al-Saud regime. In this step, we penetrated a system of Aramco company by using the hacked systems in several countries and then sended a malicious virus to destroy thirty thousand computers networked in this company. The destruction operations began on Wednesday, Aug 15, 2012 at 11:08 AM (Local time in Saudi Arabia) and will be completed within a few hours.
>
> This is a warning to the tyrants of this country and other countries that support such criminal disasters with injustice and oppression. We invite all anti-tyranny hacker groups all over the world to join this movement. We want them to support this movement by designing and performing such operations, if they are against tyranny and oppression.
>
> Cutting Sword of Justice (Cutting Sword of Justice 2012)

The Shamoon virus itself, which is more technically known as W32.Disttrack, consists of three elements. One element, the ‘dropper’ is the main component of the intrusion and inserts other modules into the target. The ‘wiper’ is one of those modules and is responsible for the destructive functionality of the virus. In the attack on Aramco, the wiper deleted the main boot records of the infected computer which meant that the computer was unable to start. Another inserted module, the ‘reporter’, sends information about the success of the compromise to the attacker. This general structure is a common pattern in computer viruses.

As promised by the ‘pastie’ in Pastebin, on the 15th of August the computers at Saudi Aramco started to fail. An estimated 30,000 workstations were affected, which is approximately three out of every four at Saudi Aramco. The economic and business impacts of the hack are widely published and easily imaginable given that normal operations were not restored for at least a week and possibly longer. CNN spoke to one of those experts contracted to help with the response, Chris Kubecka, and reported,

> Managing supplies, shipping, contracts with governments and business partners -- all of that was forced to happen on paper. Without the Internet at the office, corporate email was gone. Office phones were dead. Employees wrote reports on typewriters. Contracts were passed around with interoffice mail. Lengthy, lucrative deals needing signatures were faxed one page at a time. The company temporarily stopped selling oil to domestic gas tank trucks. After 17 days, the corporation relented and started giving oil away for free to keep it flowing within Saudi Arabia. (Pagliery 2015)

As part of the response, Saudi Aramco aggressively purchased 50,000 hard drives directly from manufacturers, which in turn disrupted global supply chains. The decision was made that simply replacing hard drives was a more expedient solution than attempting to remove the malware from existing hardware.

It took five months for Saudi Aramco to bring its systems fully back online. The expense was huge and only an organisation such as Aramco could have survived. The fact that Saudi Aramco was a state-owned company is critical in its survival and the analysis of the attack. Shamoon was an act of sabotage in that its target was systems of production, but it may also be considered as a significant attack on a nation state. Given that Shamoon targeted control and production systems and created distinct physical results, it is evident that it was a politically motivated example of cyber sabotage.

That attack did not stop with the attack on Saudi Aramco but provides a good example of the persistent nature of cyber conflict. In 2017, Cyber security company, McAfee, published a report titled, “Shamoon Returns, Bigger and Badder.” (Samani & Beek 2017) The report documents major campaigns that involved descendent versions of the Shamoon virus that occurred in 2016 and 2017 and targeted Saudi Arabia. Key factors in the analysis are that much of the code of the Shamoon virus was reused - suggesting that this was a continuation of the same campaign - and that the changes had made the effects of the malware even more serious. In particular, an updated version of the ‘wiper’ has made recovery of data impossible. In the original attack, data from the hard drives could be recovered before they were replaced, which allowed the process of hardware replacement. This is not possible with the more evolved versions of the malware.

The fact that the attack on Saudi Aramco was persistent in nature is worth keeping in mind when considering the next example of cyber sabotage which is usually considered a single ‘stand-alone’ attack. This conception is most likely misleading.

The first reports of a computer virus that became known as Stuxnet appeared in 2010. Stuxnet is an extremely sophisticated computer worm - a type of computer malware that replicates itself without human intervention and spreads by copying itself from computer to computer. When it was first discovered it was evident that it was an unusually complex worm. Liam O'Murchu, from Symantec, was on the team that first investigated Stuxnet and is reported (Fruhlinger 2017) as saying that Stuxnet was "by far the most complex piece of code that we've looked at — in a completely different league from anything we’d ever seen before." Since that point, it has probably become the most analysed and discussed example of a cyber-attack to date.

What made Stuxnet particularly interesting was that it targeted very specific industrial control processes, namely Siemens programmable logic controllers (PLCs). PLCs are at the heart of many industrial processes and provide local control over elements of the overall system. The fact that the target was PLCs meant that although this was a cyber-attack, its intention was to create damage in the physical domain. This fact explains why Stuxnet is such an important example as it emphasises the

transversality of cyber-attacks and denies the thought that cyber actions are only important in IT and computing terms.

The initial investigation of the malware rapidly discovered that PLCs were the target. What took a bit longer to parse was the fact that a very specific configuration of PLCs was targeted. As O'Murchu explains,

> We could see in the code that it was looking for eight or ten arrays of 168 frequency converters each. You can read the International Atomic Energy Association's documentation online about how to inspect a uranium enrichment facility, and in that documentation they specify exactly what you would see in the uranium facility — how many frequency converters there will be, how many centrifuges there would be. They would be arranged in eight arrays and that there would be 168 centrifuges in each array. That's exactly what we were seeing in the code. (Fruhlinger 2017)

At this point, because of the nature of the target, the complexity of the coding, and the lack of financial gain, it was obvious that this was the work of an international operation by a nation state rather than an individual hacker or group of hackers. When Symantec released this information in 2010, international analysts had been aware of problems that the Iranians had been having with their centrifuges at the Natanz processing plant, and the narrative was rapidly pieced together.

What is widely held to be true is that the United States, with the help of Israel, undertook a long-term campaign to sabotage the development of the Iranian nuclear processing capabilities. The campaign, which had the code name 'Olympic Games', likely started much earlier than 2010, perhaps as early as 2005 (e.g., Finkle 2013), and was intended to be covert. Various iterations of the Stuxnet worm were deployed with significant differences that are discussed in the following sections, and the discovery of the malware was only due to an error in one of these versions.

Stuxnet was an example of politically motivated cyber sabotage. It targeted control and production systems and had dramatic effects in the physical domain. How effective was the campaign of sabotage? Absolute figures will probably never be known, but over a period of years, the campaign destroyed something like 1000 of the 5000 centrifuges used to purify Uranium and created a level of fear, uncertainty, and mistrust in the Iranian military program (Albright et al. 2010). Some estimates say that that program was delayed by two years or more although these figures are actively contested. Iran acknowledged that there was an attack on its nuclear facilities and there had been some disruption. President Mahmoud Ahmadinejad is quoted (Reuters 2010) as saying, "They succeeded in creating problems for a limited number of our centrifuges." The importance of Stuxnet as an example of cyber conflict is not in the extent of the damage created but, rather, the fact that physical damage was created in the infrastructure of a nation state by a hostile nation state by cyber methods. Many analysts take this as proof of concept.

Stuxnet was a long-term campaign that undoubtedly involved the work of an extended team of coders. Recent analysis suggests that the contributors were more diverse than originally anticipated. For instance, in a 2019 conference, it was reported,

> ... in addition to the APT groups already linked to Stuxnet, including developers behind Duqu, Flame and the NSA-linked Equation Group, a fourth, previously unknown collaborator called Flowershop is also related. (Seals 2019)

Given this fact, it is not surprising that Stuxnet has spawned a family of related malware. 'Duqu' uses the technology to log keystrokes and mine data from industrial facilities. 'Flame' was spyware that recorded Skype conversations, logged keystrokes, and collected keystrokes. 'Havex' gathered industrial information, particularly in the aviation sector. 'Industroyer' was aimed at power companies and was the cause of blackouts in Ukraine in 2016. Triton targeted safety systems in a petrochemical plant in 2017 (McAfee 2019). Although Stuxnet in its initial form is probably a thing of the past, its successors which are all based on its technology and the lessons learned from the Stuxnet campaign continue to thrive.

The three examples presented here, Maroochy Shire, Shamoon, and Stuxnet have significant differences. They do share common factors that make them acts of cyber sabotage. In all cases, cyber methods were used to target systems of control and production. In all cases, direct physical results were achieved.

Having said that, there are subtly different interpretations of these events that are available. Intuitively, Stuxnet is usually taken as a clear example of cyber sabotage and this corresponds to the fact that the attack on systems is clear, and the intended results are evidently largely physical. This supports the use of the definition of sabotage that has been defended in this chapter. Given that, it would seem equally clear that Maroochy Shire was cyber sabotage.

It might be asked if this is simply an example of vandalism rather than sabotage, but this is misunderstanding the categorisation provided by the Influence Table. The act is sabotage because it uses systems influence to produce a physical result. If that act's only aim is to wilfully destroy or deface then it may well be also an act of vandalism. If there is some other purpose such as a political purpose it might be that it is not only vandalism. There is no reason the categorisations need to be hard-edged or exclusive.

The understanding that categorisation is not exclusive is particularly important when applied to Shamoon. It might be argued that the effects were as much informational as physical. What stopped the correct functioning of Saudi Aramco was lack of access to its data. Conversely, the response of Aramco - to simply replace the hard drives - means that the effects are regarded as physical. On a broader scale, the effects were to interrupt the production of oil and in this scope the effects are physical.

INFLUENCE TABLE	Direct Influence	Systems Influence
Physical Effects		Sabotage (Discrimination)
Informational Effects	Espionage (Proportionality)	Subversion (Just Cause)

What this means for the analysis of Shamoon is that there is some ambiguity as to whether it should be categorised as sabotage or subversion. If we regard it as an attack on the oil-producing systems of Saudi Aramco, then it is best categorised as sabotage. If we regard it as an attempt to undermine the organisational structure of Saudi Aramco, then, as is discussed in Chapter 6, it might be best categorised as an act of subversion. The lesson from Maroochy Shire is valuable here - that categorisation is not hard-edged or exclusive. The Shamoon attack can be regarded as both an act of sabotage and an act of subversion because it has effects both in the physical and informational domains.

The definition of cyber sabotage that was proposed was that it is the exertion of systems influence by cyber means with the aim of creating physical effects. The diverse examples provided in this section support that definition and elucidate it in several ways. Firstly, the categorisation is inherently neither hard-edged nor exclusive. Secondly, some of the characteristics of cyber methods have emerged in the discussion, again confirming earlier theories. Cyber sabotage can have dispersed effects - the effects may be experienced by large numbers of the population rather than being tightly focused on particular individuals. Cyber sabotage campaigns can be persistent - the rivalry that a cyber sabotage campaign embodies is likely to be continuous or ongoing rather than a single one-off event. Cyber sabotage campaigns may be covert in the broad sense of the word in that they can be non-overt and easily deniable.

4.4. Epistemic Problem of Recognition for Discrimination

This problem of knowing who is liable to harm is an epistemic problem. It is worth separating it from the conceptual problem of describing or delineating who is liable. For instance, there is some uncertainty about what constitutes a non-combatant. For instance, do workers manufacturing weapons qualify as combatants or not? These are thorny issues regarding delineation, but they are not the focus here. The focus is

that, given any resolution of these issues, does one possess the information to apply the understanding of that resolution correctly? So, even if we have a conceptual definition of who is a combatant and who is not, do we possess the information and skills to correctly apply this definition? The problem exists in any model where some people are conceptually held to be liable while others are not.

There are two reasons why this epistemic problem is aggravated in the case of cyber conflict. The first is the covert nature of cyber conflict. The second is the distancing of cyber techniques from the conventional military.

4.4.1. The Covert Nature of Cyber Methods

One of the lessons from Stuxnet was that with enough forensic investigation a cyber-attack might be correctly attributed. As discussed earlier, the application of ‘covert’ to cyber methods is sometimes problematic and it often appears that ‘deniable’ would be the more appropriate concept. However, the characteristic - which I will call its 'covert nature’ - is meaningful in this context. Attribution is usually possible but is often not immediate and often not entirely assured. It may be that the attribution satisfies the demands of reasonable doubt rather than those of absolute certainty. Even in this broad context, the covert nature of cyber methods poses problems for discrimination.

The covert nature of cyber methods results in a state of affairs in which we are uncertain who exactly is responsible for a particular threat. Regardless of what one’s standpoint is regarding the technical definition of liability, any ignorance or uncertainty regarding the responsibility for threat is devastating to the application of the Principle of Discrimination.

The Saudi Aramco example provides strong support for this statement. The attack was announced prior to the activation of the malware. Responsibility was overtly accepted by the ‘Cutting Sword of Justice’ group. The political motivation of the attack was overt. Yet the precise individuals who undertook the attack remain unknown - at least publicly. We may never know if the intelligence agencies involved have greater information but let us assume that at the time no greater knowledge existed. At the point when the threat was received but the attack had not yet caused damage how might we apply the Principle of Discrimination? At this point, some counteraction might be perceived as warranted and proportionate. The fact that Saudi Aramco and the Saudi government were unaware of exactly who was responsible was an insurmountable problem in responding. It was a failure to meet the Principle of Discrimination which prevented any counteraction from being possible. This is a failure of Discrimination of the most brute type, where one has no idea who is responsible, and therefore no idea of who is liable.

Not all failures of discrimination are so blunt. Imagine for instance a country (A) has experienced an extensive and ongoing cyber campaign against its power and electricity infrastructure that shows no sign of letting up. The country is involved in a long-term rivalry with another country (B) and the evidence is sufficient to believe that

B is the source of the campaign. Some response to the threat that is posed by (B) may seem reasonable. But meeting the requirements of the principle of discrimination is made difficult in cases where one cannot be sure of those who bear responsibility. Tough questions arise. For example, would it be permissible to target a group of ten programmers even if one knew that only seven of them were responsible and one did not know which seven? The complexity of real-world examples such as these is created by the interaction of epistemic uncertainty and the vagueness surrounding the ideas of group responsibility versus individual responsibility and does not have easy answers.

The problem that the covert nature of cyber methods creates is that one needs to be assured that one is targeting the correct people. The stakes are high because significant harm may be allowable in certain circumstances if the threat that is to be avoided is significant. This is a problem that is not unique to cyber methods. Other covert actions pose the same challenges. However, a covert nature is central to the nature of cyber methods, whereas it is more exceptional in terms of conventional actions. A conventional action will be overt unless there are specific efforts to make it covert. A cyber action will be non-overt and deniable unless the actor makes specific efforts for it to be overt.

4.4.2. The Distancing of Cyber from Conventional Military

As was mentioned earlier, Jason Healey is best known for the excellent discussion of cyber war presented in *A Fierce Domain,* but he was previously instrumental in forming a scale of responsibility for cyber actions. This process started at the Joint Task Force for Computer Network Defense, and has now become the "Spectrum of State Responsibility for Cyberattacks" which is published by the Atlantic Council (Healey 2011 & 2013). In outline form it is:

1. State prohibited
2. State prohibited but inadequate
3. State ignored
4. State encouraged
5. State shaped
6. State co-ordinated
7. State ordered
8. State rogue executed
9. State executed
10. State integrated

The spectrum provides a useful tool for analysis. In the context here, it is used to demonstrate the range of involvement of state authority in cyber actions. Traditionally, the harm creating actions involved in large-scale conflict have been predominantly the responsibility of state militaries. Of course, states have used some irregular forces. Also, some conflicts have involved groups other than states. However, in the

case of large-scale cyber conflict the substantive difference is that very few of the actions are conducted by the military. Even towards the bottom of the list - those actions that are directly sanctioned and implemented by the state - the actors involved are more likely to be members of the intelligence agencies or other agencies that are not directly associated with the state military.

A significant symptom of this state of affairs is that the most prolific and dangerous cyber operations are referred to as advanced persistent threats, APTs. A non-technical list (Mandiant 2021) of APTs is maintained by the security company FireEye which describes it as a “Who's who of cyber threat actors.” A typical entry is APT41, “a prolific cyber threat group that carries out Chinese state-sponsored espionage activity in addition to financially motivated activity potentially outside of state control.” There is undoubtedly a degree of cultural bias in the FireEyes list. ATPs are listed in Iran, China, Russia, North Korea, and Vietnam. None of the listed APTs is in the U.S. or the allies of the U.S..

The entry for APT30 is also relevant, “Evidence shows that the group prioritizes targets, most likely works in shifts in a collaborative environment and builds malware from a coherent development plan.” The threats are ‘advanced’ and ‘persistent’ which means that they are based on more than a loose collective of programmers. What is known about them is based on supposition as much as evidence. The claim of a shift working pattern is usually based on the timing of actions - in the same way as evidence for Russian actions involving the U.S. elections was narrowed down because the activities all occurred in the working day of Moscow time (Mandiant 2021).

It is easy to get embroiled in the stories of APTs that include players with names such as ‘Tsar Team’ and ‘Deputy Dog’, campaign names such as ‘Clandestine Fox’ and tools such as ‘Ghost Rat’, ‘Gothic Panda’ and ‘Bug Juice.’ What is important here is the way in which this form of operation causes problems for discrimination. The Spectrum of State Responsibility outlines the various relationships that are possible between the state and the actual perpetrators of cyber-attacks. The perpetrators of harm-inducing actions are no longer likely to be only the military. The range of actors and their relationship with the state obscures the certainty of attribution that is necessary for the Principle of Discrimination.

4.4.3. A Case Study

APT3 is one of the oldest and most sophisticated threats. It is known for a combination of phishing attacks and the use of zero-day exploits. Phishing attacks are low tech in that they usually consist of an email, often sent in bulk to the target’s employees, that contains a link that, if clicked, will lead to some seemingly innocuous web page that will additionally attempt to install malware on the user’s computer. Zero-day exploits are high-tech and are unique vulnerabilities that are discovered in computer operating systems and other software. Zero-day exploits are precious because once they are discovered they can be patched - the vulnerability can be rectified. APT3 is used as

an example here because there is a high degree of certainty that it is connected to China's Ministry of State Security (MSS) despite operating from the Guangzhou Boyu Information Technology Company (Insikt 2017). APT3 provides clear demonstration of both causes of the epistemic problem of recognition described above.

> APT3 is a China-based threat group that researchers have attributed to China's Ministry of State Security. This group is responsible for the campaigns known as Operation Clandestine Fox, Operation Clandestine Wolf, and Operation Double Tap. (Mitre Att&ck 2017)

As an example, Operation Clandestine Fox demonstrates the covert nature of cyber methods. The aim of the operation was to install backdoors in corporate systems. As FireEye reports, APT3 were "extremely proficient at enumerating and moving laterally to maintain their access" (Eng and Caselden 2015). The original compromise, which was aimed at high-tech U.S. industries such as aerospace and defence, relied on both low-tech and high-tech methods. Email phishing was used to get employees to click on a link with the promise of, for example, savings on the price of an Apple computer. The link leveraged a sophisticated zero-day exploit of video files to enable installation of further code that would allow future access to the system.

It is relatively simple to obscure the source of phishing emails. Likewise, it is relatively simple to hide the location of links that are to be clicked on. In order to trace the source of an attack of this nature, high-level forensic techniques are needed. Those techniques themselves are covert as they form part of a nation's security structure. Moreover, analysis might be made easier if the compromise was left in place and computer activity on the compromised machines could be monitored. This is rarely acceptable. Certain attribution of an attack like this is impossible in the short term. In the medium to long-term, greater certainty is certainly possible. Only in one case - Stuxnet - has there been such a weight of evidence that some of the responsible actors were forced to acknowledge their complicity. In all other cases, the responsible actors have felt that denial was a credible and plausible policy. Although it may be strictly incorrect to describe cyber methods as covert, they are always deniable.

The example of APT3 also demonstrates the distancing of cyber from conventional militaries. If we accept the assessment that APT3 is closely related to China's Ministry of State, this means that it is not overtly a military operation. On the Spectrum of State Responsibility, it is likely that APT3 sits somewhere in the region of level 4 (state encourages) to level 7 (state ordered). Proving that it warranted a higher level than that would be difficult. It seems likely that the Chinese authorities have set up APT3 specifically to separate its activities from conventional military operations. The example of patriotic hacking described in Chapter 1 - in which hackers such as 'Black Hand' intervened by targeting NATO systems 'on behalf' of their cause - provides an example in which the responsibility is further devolved. The reason the Spectrum of State Responsibility is a powerful tool is that it both acknowledges and quantifies the ways a state can be involved. How actors position themselves on the spectrum is often determined by their tolerance to attribution. Regardless of this, as we move

away from the ‘State Executed’ and ‘State Integrated’ levels of responsibility, the problems for discrimination become greater.

The fact of the matter is that governments and other organisations can leverage the uncertainties described here. By distancing the operations from the conventional military, they can increase the obstacles of rivals meeting the requirements of the Principle of Discrimination, thereby making counteractions more difficult to justify. It is in the interest of states to separate their offensive cyber operations from the military and obscure the operating structure. In the world of long-term geopolitical rivalries, it seems entirely implausible that all those involved are not doing exactly this. That is not to say that the associated militaries do not have cyber capabilities, only that in the case of offensive cyber capabilities there is an advantage in separating those operations from the military.

4.4.4. Inadequate Epistemic Certainty

In this section, the epistemic problems of recognition that affect the Principle of Discrimination have been discussed. The major epistemic problem of recognition, of being able to determine who is liable, is driven by two factors. Firstly, the covert nature of cyber methods means that certainty of attribution is unlikely. Secondly, the structure of the organisations responsible for offensive operations is largely opaque and it is in the interest of states to keep them so, precisely to muddy the waters for the Principle of Discrimination. This second face of the epistemic problem is shown by the distancing of offensive operations from the conventional military.

The two problems discussed here are intrinsic to cyber methods. Cyber methods have a covert nature and cyber methods are not necessarily associated with the military. This is a substantive difference to conventional methods which are not inherently covert and are usually associated with the military.

With regard to cyber conflict, those wishing to defend the use of the Principle of Discrimination may point out that similar epistemic problems exist in conventional conflict. We do not always know which forces are responsible for a particular attack, it is not always an established military force and attacks may be contracted out to freelancers. In response to this defence, one might comment that this ‘outsourcing’ of conflict is indeed not unique to cyber conflict, but perhaps there is a substantive difference in the fact that it is the main mode of operation in cyber conflict. More important than that is the fact that my purpose is to undermine that justification of the application of the Principle of Discrimination in the context of cyber conflict. Saying that it is potentially equally undermined in other contexts does not provide an obstacle for that purpose.

4.5. Epistemic Problem of Accuracy for Discrimination

In the aftermath of Stuxnet, when it became apparent that there had been the first known truly major cyber-attack with physical effects a particular phrase from Ralph Langner was widely quoted (Broad et al. 2011). "The attackers took great care to make sure that only their designated targets were hit. It was a marksman's job", said Langner at the time. Langner is often credited for 'discovering' Stuxnet and devoted a great deal of time to forensic analysis, becoming an accepted authority on the attack.

It is tempting to believe that that single phrase has influenced the subsequent discussion more than it should. Whether or not that is true, what is certainly the case is that there is a theme in the discussion that argues that cyber methods are capable of superior discrimination than conventional methods.

The epistemic problem of accuracy for discrimination suggests that even if we know the people that we are attempting to target, there may be significant uncertainty about the ability of our methods to inflict harm only on those people. The counterclaim is that cyber methods are particularly discriminate in this respect. One of the most cogent defences of the claim is provided by Ryan Jenkins' chapter (Jenkins 2016), "Cyberwarfare as Ideal War." Jenkins, talking about cyber methods, says,

> A war could maximally appropriately distribute any harms that did occur by ensuring that only those liable to attack were actually harmed. (Jenkins 2016:95)

Are there reasons to believe that cyberwar is more able to be discriminate than conventional war? As we will see, Jenkins believes that there are. The claim that cyber methods are capable of superior discrimination than conventional methods is a comparative one and may be interpreted in a number of ways. The first interpretation is that cyber methods are inherently more able to approach ideal discrimination than conventional methods (termed here *inherent superiority*). Alternatively, perhaps the best cases of cyberwarfare are more able to approach this ideal than the best cases of conventional methods (*maximal superiority*). In addition, there is a claim that cyber methods are in practice more able to meet these ideals (*empirical superiority*), although given the limited experience we have of cyber conflict such a claim is open to a fundamental challenge of remaining unproven. The claim must be defensible for at least one of these interpretations.

These categorisations of superiority are useful as we move through an analysis of Jenkins' defence of the claim that cyber methods are superior to conventional methods - at least potentially. His argument proceeds in two parts. Firstly, a fictional example is used, in an exploratory manner, to show that in this idealised case ideal cyberwarfare is possible. Secondly, a real-world example, the Stuxnet worm, is used to show that the fictional example is not so far-fetched as to be meaningless.

Cyber 1

> A militarily weak state, R, shares a border with its much more powerful neighbour S. In a moment of geopolitical avarice, S invades R. Suppose that repelling S's invasion of R would be a just cause for war, but because S has a powerful conventional military, the disvalue caused by the disproportionate harms R could expect a conventional war with S to cause would easily eclipse the value to be won by repelling the invasion. Now, T is a state with sophisticated cybercapabilities and a strong commitment to international humanitarian law. T appreciates that it could leverage its cybercapabilities against S in order to coerce S to withdraw from R. Accordingly, T offers S an ultimatum to withdraw from R or else suffer serious cyber consequences. S refuses, and T responds with a cyberattack. S's conventional military capabilities are disabled—its command and control centers become inoperable, its logistics databases become scrambled, its lists of targets in R (and T) are encrypted and rendered inaccessible, and so on. Furthermore, T's cyberattack is so discriminatory that only those people and services that are legitimate targets are impacted by the attack. S sues for peace, and T lifts the veil, restoring S's conventional military capabilities, on the condition that S withdraw from R. (Jenkins 2016:96)

He takes it that an intervention such as that in **Cyber 1**, is maximally discriminate. As this is a bland claim (the example states explicitly that the response is discriminatory although the method of discrimination is not described), the example of Stuxnet is resurrected in order to show that such discrimination is possible in the real world.

Almost uniquely, the level of documentation and forensic analysis of Stuxnet forced the perpetrators, in this case, the U.S. and Israel, to acknowledge their complicity. More is known about Stuxnet than about almost any other cyber action. Jenkins' claim is that Stuxnet was written specifically to target the Iranian Natanz facility. He notes that the Stuxnet worm has limited abilities to propagate itself across networks and that there is no known case of Stuxnet harming other installations. This mirrors the earlier quote of Ralph Langner (Broad et al. 2011), "The attackers took great care to make sure that only their designated targets were hit. It was a marksman's job." This is enough, Jenkins claims (Jenkins 2016:97), to "give us confidence about the capacity for cyberweapons to inflict discriminate harm."

It is not enough to give us that confidence for several reasons. Firstly, one example would never provide absolute confidence. Secondly, the details of Stuxnet remain at least partially opaque. It was a covert operation and the narrative that has developed around it is not entirely consistent. Thirdly, it seems highly likely that Stuxnet did not perform exactly as intended. Perfect discrimination may have been intended but it was not achieved.

Even though the U.S. has acknowledged complicity in the attack, full details of the Stuxnet worm itself remain contested.[21] The details of the project that created and deployed it, apparently named 'Olympic Games', are equally opaque. Despite these facts, an accepted narrative has developed around Stuxnet. Some elements of that narrative are at best puzzling, and at worst inconsistent. One element of the narrative that is usually emphasised is the fact that the systems at the Natanz facility were 'air-gapped', in other words, disconnected from other networks and the Internet. The challenge then was how to get the worm into the facility's systems. The method of infiltration of the first version of Stuxnet is uncertain but likely involved USB drives. The method of infiltration of subsequent versions of Stuxnet is more clearly documented. The virus was infiltrated into the systems of contractors who did work at the Natanz facilities in the hope that the virus could be carried into Natanz by their equipment (Zetter 2014). The existence of multiple versions of the worm and the varied capabilities of those versions and the alternative methods of infiltration are usually not present in the popular narrative.

The presence of multiple versions of the Stuxnet worms causes problems for the claim that it was discriminate. Firstly, it is documented (Zetter 2014 & Mimoso 2014) that users at one of the contract companies, Neda, reported issues outside the Natanz facility with software associated with the Siemens Step 7 PLC that was the target of Stuxnet. Not all the issues were contained within the target. Secondly, the claims that are prevalent about the inability of Stuxnet to spread are simply untrue. Stuxnet spread aggressively. In 2010, Thomas Chen (2010) estimated that between 50,000 and 100,000 computers had been infected. Although earlier versions of Stuxnet had limited abilities to spread, later versions had much greater abilities (e.g., Falliere et al. 2010),

> To ensure greater success at getting the code where it needed to go, this version of Stuxnet had two more ways to spread than the previous one. Stuxnet 0.5 could spread only by infecting Step 7 project files—the files used to program Siemens PLCs. This version, however, could spread via USB flash drives using the Windows Autorun feature or through a victim's local network using the print-spooler zero-day exploit that Kaspersky Lab, the antivirus firm based in Russia, and Symantec later found in the code. (Zetter 2014)

Ralph Langner notes that the aggressive spread of the second version of Stuxnet is due not to direct connection to the Internet.

> All of a sudden, Stuxnet has made its way around the globe — not because of the fact that billions of systems are connected to the Internet, but because of the trusted network connections that tunnel through the Internet these days.

---

[21] For instance, Liam O'Murchu at Symantec, is reported as saying that the code was never released and could not have been reverse engineered despite numerous claims to the contrary. (Fruhlinger 2016) Much of the mythology about Stuxnet is constructed on limited information.

> For example, remote maintenance access often includes the capability to access shared folders online, giving Stuxnet a chance to traverse through a secure digital tunnel. My colleagues and I saw exactly that when we helped Stuxnet-infected clients in industries completely unrelated to the nuclear field back in 2010. (Langner 2013)

It is simply incorrect to claim that Stuxnet was a neatly targeted worm that only appeared in the target systems. It spread widely, was difficult to contain and was originally discovered by chance on systems that were entirely remote from the target systems. Moreover, it seems likely that its spread was, if not intentional, anticipated. Ralph Langner's interpretation is illuminating,

> Given that Stuxnet reported Internet protocol addresses and hostnames of infected systems back to its command-and-control servers, it appears that the attackers were clearly anticipating (and accepting) a spread to noncombatant systems and were quite eager to monitor that spread closely. (Langner 2013)

There remains the claim that although Stuxnet did, in fact, escape from its intended target network it was ineffectual outside that network. It is true that Stuxnet seems to have been created to search out very specific system configurations and only become 'active' in those circumstances. This claim, that Stuxnet was not harmful outside the target area has two interpretations. The first is that there was, in fact, no harm done by the Stuxnet virus outside the target systems. The second is that the infiltration of software onto systems, while it may be wrong, is not harmful until that software becomes active or is used in some way.

Whether there was harm caused directly by the Stuxnet worm outside the target systems is difficult to determine. Anecdotal evidence, such as the evidence of the forum posts by a control engineer working for Neda that they were having problems with equipment is not unassailable proof. Moreover, issues around the world that were initially ascribed to Stuxnet have since largely been attributed elsewhere (Pauli 2013 & Carr 2010). However, major security companies, Kaspersky and Symantec were instrumental in the response to Stuxnet and in its report Symantec states,

> While their choice of using self-replication methods may have been necessary to ensure they'd find a suitable Field PG, they also caused noticeable collateral damage by infecting machines outside the target organization. The attackers may have considered the collateral damage a necessity in order to effectively reach the intended target. (Falliere et al. 2010)

It seems likely that there was at least some collateral damage associated with Stuxnet. Certainly, extensive expense and resources have been used in an attempt to counter that damage.

However, the definition of what amounts to collateral damage as used by Symantec may vary from common usage. In my personal experience, for computer security experts the infiltration of unauthorised software into a system would always constitute

both a wrong and a harm. I take it that the wrong is intuitive and needs no defence. Perhaps, though, in the case where the software has no effect, such an act amounts to a 'harmless wrong' of the type most fully discussed by Joel Feinberg and Arthur Ripstein (Feinberg 1990, Ripstein 2006). It is beyond the scope here to complete the discussion of whether malware in and of itself is harmful. In technical terms, the worm included both a user-mode and kernel-mode rootkits. In plainer terms, this means that they subvert security at a deep level in the systems. This, for security professionals, is the worst type of system compromise and invalidates many of the security measures that have been put in place. System security is therefore damaged.

The cyber interests framework developed in Part One emphasised two important interests; private data control and public data access. Those cyber interests have played a lesser role in this chapter because sabotage has its effects predominantly in the physical domain. However, a compromise of system security such as the installation of unauthorised rootkits is a major degradation of private data control. Therefore, in terms of this project, it can never be seen as harmless.

Claiming that the unauthorised installation of rootkit malware is not harmful is equivalent to claiming that the installation of unmonitored doorways to military installations would not be harmful until those doorways were actually used. There may be a philosophical argument that would support this conclusion. It is not pragmatically meaningful because most of us would think that those doorways created a harmful security risk. Likewise, it is hard to see that installation of rootkit malware is not harmful just because the rootkits were possibly not used. Given this, it is hard to see how the claim that no harm was done outside the Natanz facility is tenable. The worm spread to computers and machinery all over the world. Considerable expenditure was dedicated to fighting the results of this spread. Even if there was no direct physical harm comparable to that at Natanz, which is not proven, there was considerable other harm. The legacy of that harm continues in the family of attacks that Stuxnet has spawned. It would not be tenable to say that Stuxnet demonstrates how cyber methods have *empirical superiority.*

As Stuxnet does not provide an example of *empirical superiority* in terms of discrimination that supports Jenkins' claim then perhaps they are *maximally or inherently superior* to conventional weapons?

A claim of inherent superiority would imply a commitment to the idea that the harms of cyber methods were inherently less than those of conventional methods. Given the transversality of cyber methods, this seems implausible. One way this claim has sometimes been framed is based on the idea that cyber methods only target IT systems, and those systems can never be as valuable as, for instance, human lives. There are two reasons why this idea is misleading. The first is the transversality of cyber methods. Cyber methods can have effects in the physical domain. The second is that cyber interests can have significant value in our lives as demonstrated in Part One as part of the development of the cyber interests framework. A cyber method that degrades any of the capabilities that are necessary for a meaningful human life

is significant. In this context, it seems impossible to provide a rigorous defence of the inherent superiority of cyber methods.

Maximal superiority is equally difficult to defend. The thought here is that discrimination can be incorporated in the method at the design stage whereas that is impossible with a conventional weapon such as a rifle. This argument is flawed because it is not the weapon that is either discriminate or not discriminate. It is its use. The sniper is able to discriminate his target. The tank commander is able to discriminate his target. The drone operator can certainly discriminate their target. In the case of the sniper, almost perfect discrimination is possible. For cyber methods to be maximally superior, it would need to be demonstrated that their level of discrimination could be better than this.

That is not to say that there are no conventional weapons that lack this ability to discriminate - many missiles, bombs, and such are indiscriminate. But equally, I do not deny that there are cyberweapons that are indiscriminate. In order to support the idea of maximal superiority, it is necessary to compare the best cases of both sides. In general, it seems that the best cases of cyber methods are often compared to the generic performance - admittedly horrific - of conventional methods. Ultimately, it seems that the discrimination of an act of war is dictated as much by the operational constraints provided practically and politically, as by the choice between cyber and conventional weapons.

Jenkins provides little support for the claim of discrimination other than the example of Stuxnet and it remains unclear that cyberweapons are either *maximally or inherently superior* to conventional weapons. Worse for the advocates of this claim, there are some reasons to believe that the opposite is true; that there are reasons that cyberweapons are likely to be less discriminate.

Cyberweapons, for the most part, rely on vulnerabilities in the target system.[22] For an attack such as described in **Cyber 1** to be successful, vulnerabilities would have to exist in S's systems. The fact that T has 'sophisticated cyber capabilities' does not mean that they could successfully attack a system without vulnerabilities. Cyber-attacks rely on vulnerabilities, and this is a substantive difference to conventional weaponry. A bomb will do some damage regardless of the defence. On the other hand, a defended computer system, isolated from the internet, is extremely difficult to compromise. Put simply, an individual tank is a hard target for a hacker.

Military systems are isolated and hardened against intrusion. That is not to say that they cannot be compromised. Such compromises have been demonstrated in the real world (e.g., Weinberger 2007). However, a cyber-attack of that sort is likely to be a 'single use' deployment. It will rely on a vulnerability that is unknown to the defenders. Its use by the attackers is likely to reveal it to the defenders who will be able to protect

---

[22] Again, the exception is denial of service type attacks which rely on overloading the capabilities of systems.

against further compromises by this vector. It is probable that military systems will become increasingly hard to compromise. In comparison to other systems, they have two advantages. They can be isolated from the Internet and other communication, and they have a degree of top-down control which allows security measures to be enforced.

In contrast, the general, publicly accessible Internet does not have those advantages. There is extremely little top-down control, and the general implementation of security measures relies on the actions of individuals. Both this and the inherent connectivity are aspects of the Internet that are protected as essential to personal liberties. What that means is that security compromises and breaches on the Internet are much more plausible than on military networks. This results in the fact that an easy way of applying political pressure to a nation is to apply very widespread pressures to that nation's citizens. This is the form, closely related to that which is known as hacktivism, which is likely to become increasingly prevalent in cyber conflict and is in direct opposition to discrimination. George Lucas defines hacktivism as malevolence in the cyber domain which may be an apt description (Lucas 2017:20-22). Though the term hacktivism suggests counter-authority actions by small groups, the technique is available to larger groups such as nation states. The fact that highly discriminate attacks are likely to become increasingly difficult means that highly indiscriminate attacks are increasingly likely. This corresponds directly to the idea that cyber conflict is characterised by dispersed action and effects.

The epistemic problem of accuracy for discrimination with respect to cyber methods reflects the idea that cyber methods are not certain to target those at whom they are aimed. There are no persuasive reasons to believe that cyber methods are either maximally superior or inherently superior to conventional methods. Likewise, there are no persuasive examples that demonstrate empirical superiority in discrimination over conventional methods. Therefore, one cannot be assured that cyber methods are in any way superior in terms of discrimination than conventional methods. Moreover, there are persuasive reasons to believe that the opposite is true.

4.6. Conclusion

The principle of discrimination demarcates those people who are permissible targets for the imposition of harm. There are widespread philosophical discussions of the exact configuration of the criteria that should be included in the principle. This is the conceptual problem. There are also epistemic problems with the implementation of the principle. Even if we agreed a defined principle of discrimination, we would need to be assured that we were able to separate those who match those criteria from those who do not. This is what I have termed the epistemic problem of recognition for discrimination. Secondly, we must be assured that our actions are capable of targeting that group while not producing collateral damage. This is the epistemic problem of accuracy for discrimination. In the context of cyber conflict, neither of these problems is resolved.

A successful principle of discrimination must both be conceptually coherent and capable of implementation. This demand is due to the fact that we require a pragmatic and functional principle that is action-guiding rather than simply a work of conceptual elegance. For this requirement to be fulfilled both sets of epistemic problems would need to be overcome. In fact, as has been shown, there is no reason to believe that the epistemic problems are not significant. Conversely, there are strong reasons to believe that the epistemic problems are an inherent characteristic of cyber methods.

The epistemic requirements placed on the principle of discrimination are not limited to cyber conflict. They exist in all forms of conflict. However, whereas for conventional conflict they are marginal to the methods, in the case of cyber methods they are central. In the case of cyber conflict, it is reasonable to believe that the epistemic problems take on such significance that they make a useful principle of discrimination implausible.

Readers will notice that there has been little mention of cyber interests or a cyber interests framework in this chapter. This is despite the fact that Part One was given over to defending such a framework. This chapter has largely talked about cyber methods rather than cyber interests. This forms what might be regarded as a traditional stance towards cyber conflict and is appropriate in the context of sabotage.

What is possible to say at this early point of Part Two is that a principle of discrimination is severely challenged with respect to cyber methods. Moderate theories of war rely on some theory of discrimination and therefore moderate theories of war are severely challenged with respect to cyber methods.

This conclusion is possible without a commitment to the cyber interests framework developed in Part One. There are corresponding arguments in the arenas of espionage and subversion that impact on discrimination but focusing on sabotage has allowed moderate theories of war to be challenged without a commitment to the cyber interests framework. Already, things are going badly for moderate theories of war. The next chapter will bring the cyber interests framework into the centre of the discussion and things will go from bad to worse.

5. Espionage and Proportionality

Secret operations are essential in war; upon them the army relies to make its every move. An army without secret agents is exactly like a man without eyes or ears.

Sun Tzu, The Art of War

5.1. Introduction

The previous chapter paired sabotage with discrimination as a method of highlighting issues around actions that have system influence in the physical domain. This chapter considers actions that have a direct effect on informational assets, and it pairs espionage with the Principle of Proportionality. Its aim is to demonstrate that proportionality cannot be applied in a principled manner in the context of cyber conflict.

INFLUENCE TABLE	Direct Influence	Systems Influence
Physical Effects		Sabotage (Discrimination)
Informational Effects	Espionage (Proportionality)	Subversion (Just Cause)

Linking espionage with proportionality is an unusual move because espionage does not usually appear in conventional proportionality assessments at all. In fact, espionage has a somewhat difficult and fraught relationship both with the strictures of international law and with ethical frameworks.[23] However, if cyber espionage is shown

[23] While spying is not against international law, spies can be tried under domestic law. They are not regarded as combatants, and the division between spies and combatants is commonly

to have significant effects on interests that contribute to flourishing human lives, then its exclusion from proportionality is untenable.

The cyber interests framework developed in the first half of this project demonstrates the importance of cyber interests in flourishing human lives. In particular, private data control and public data access are taken as significant human interests. Espionage, and in particular cyber espionage, can have direct influence on both those interests. In other words, cyber espionage can have direct influence on significant factors in the provision of flourishing human lives.

The inclusion of cyber interests in proportionality assessments highlights and emphasises a fault line in proportionality. Proportionality is a principle that compares the harms and benefits of one outcome to those of another. This works well if the harms (and benefits) are comparable. The comparison, as we will see later, may be direct or by some other indirect technique. If the harms on either side of the comparison are genuinely incommensurable - they cannot be compared in any way - then proportionality fails.

The incommensurability problem exists for any pluralist interpretation of harm but is emphasised when the harms considered are greatly incommensurable.[24] Historically such issues have existed between physical and spiritual interests. Cyber interests are greatly incommensurable with physical interests. It is implausible to compare the loss of private data control to a physical injury. There is no direct or indirect way in which the two can be compared in a principled manner. Without a technique of principled comparison, proportionality fails.

In order to defend the claims above, some groundwork is needed. Section 5.2 takes a brief look at the Principle of Proportionality which inevitably emerges as not without its challenges and internal conflicts. Proportionality is taken as a technique that aims to prevent excessive actions by balancing the benefits of an action against the harms produced. In conventional JWT, proportionality appears as a criterion of both Jus ad

---

made by the wearing of a uniform (something that is not meaningful in cyber conflict) This is best shown in an extract from Israel's manual on the Rules of Warfare (ICRC 2022a) which states, “Spying in itself is not an action that is prohibited in itself under the rules of war, and a country sending a spy into enemy territory is not in breach of international law. As has been shown, however, a spy does not meet the conditions required of a legitimate fighter (because he is hidden among the civilian population) and is consequently not entitled to the immunity from prosecution of a prisoner-of-war.” (ICRC 2022)

[24] The word incommensurable is perhaps used here in an unusual fashion. Incommensurable in its strictest interpretation is absolute. Two things are either incommensurable or not. Throughout this section the word is used in a looser manner in which incommensurability can have a degree. Things that are very similar are very commensurable. Things that are very different in relevant ways are very incommensurable. The point made in this footnote is linguistic. A more complete defence of a similar position is provided by Hájek & Rabinowicz (2021).

Bellum and Jus in Bello. For the most part, the examples used will reference actions taken as part of an existing conflict or rivalry rather than actions taken towards the initiation of conflict - although with the persistent nature of cyber rivalry the distinction is somewhat less defined than is sometimes assumed in classical theory. Although the examples used will mainly show Jus in Bello scenarios, the logic of the argument would apply equally to Jus ad Bellum assessments.

Sections 5.3 and 5.4 look at espionage and then cyber espionage. They demonstrate that the definition of espionage as direct influence on informational assets is tenable and does not contradict general usage. It emerges that the use of the influence table highlights elements of cyber espionage that are unique. Cyber espionage allows the influence of informational assets to be active as well as passive.

The argument as to whether proportionality can accommodate incommensurable interests starts in Section 5.5 which lays out the bones of the argument. Cyber espionage influences cyber interests. Cyber interests are significant in human lives. Therefore, cyber espionage can have significant influence on human lives. Therefore, it cannot be excluded from proportionality assessments. When it is included, the incommensurability of the harms and benefits renders proportionality non-functional.

Section 5.6 outlines various ways that one might defend proportionality from these claims. One is that incommensurability is taken too seriously in the previous paragraphs. Even if things are technically incommensurable, we are still able to make valid ethical judgements. Apples and oranges are different, but we can still decide which might be better in a cake. This is termed the judgement approach. The second defence is termed the tracking approach. It relies on the fact that the value of other harms will roughly correlate with a ‘major’ harm. For instance, maimings in conventional war roughly correlate with deaths. The more deaths result from an action, the more maimings will occur. We can therefore only consider deaths and remain confident that the conclusions would be the same if we had considered maimings and deaths.

The tracking approach is quickly discounted in the context of cyber interests because informational interests clearly do not track physical interests. Section 5.7 provides strong reasons to doubt the judgement approach. An absolute proof is unlikely as the aim would be to demonstrate a negative, that judgement cannot make principled proportionality assessments. Rather, strong doubt is provided, and the onus remains with the advocate of the judgement approach to demonstrate how this doubt might be overcome.

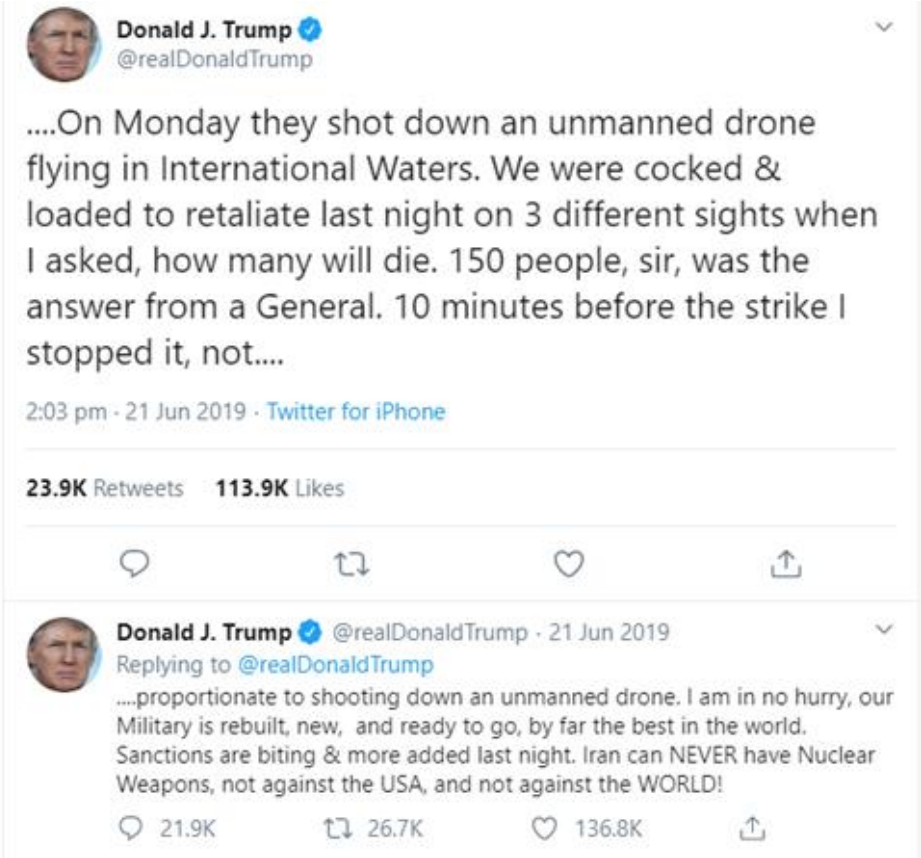
The conclusion acknowledges that it is an uncomfortable realisation that incommensurability is an insuperable problem for proportionality in the context of cyber conflict. The major question of this project is whether moderate theories of conflict can accommodate cyber conflict. If proportionality fails in this context, it seems implausible that those theories can be applied to cyber conflict in a principled manner.

The fact that this is an uncomfortable realisation naturally leads one to search for further ways in which this may be avoided. If no obvious route is found the inapplicability of proportionality combined with the problems of discrimination discussed in the previous chapter demand that moderate theories of conflict are designated as impractical in the context of cyber conflict. A fuller discussion of those issues is left for the final conclusion in Chapter 7.

5.2. Proportionality

5.2.1. The Use of 'Proportionality'

In June 2019, an American drone was shot down while flying close to the Strait of Hormuz. Donald Trump initially ordered a counterstrike against Iran and then, at the last moment, cancelled the attack because it was not proportionate.



It may not come as a huge surprise to some readers that the way Trump uses 'proportionate' is contested. Trump implies that he is balancing the harm that he intends to inflict with the harm that had previously been inflicted on U.S. forces. This

usage is based on a counter-punching model of proportionality that uses ex-ante events to justify an action.

Although this is a common use of the word, it will cause a few raised eyebrows in the academic and legal professions. In these spheres, proportionality is more commonly portrayed as a balance between the harm that is averted by one's action and the harm that is inflicted by one's action. As Adil Ahmad Haque claims, an action only satisfies proportionality if the strategic goal of the action is worthy of the harm that it inflicts.

> The importance of this prohibition is difficult to overstate. While the precautions rule regulates how armed forces may pursue a particular military advantage, the proportionality rule regulates whether a particular military advantage may be pursued or must be abandoned. Even if attacking forces select the weapons, tactics, and targets that best avoid or most reduce harm to civilians, even at their own risk, they must forego one path to victory if the expected civilian losses are too great. (Haque 2017: 88)

Of course, there is no monopoly of academia in defining the specific meaning of words, but as this section concerns a principle of proportionality to be used in specific situations in the context of large-scale cyber conflict, a certain amount of terminological scaffolding is required at this point. As Ariel Colonomos notes,

> We find proportionality as a norm in three overlapping spheres. It is codified in both international law (IL) and international humanitarian law (IHL). It is a moral norm embodied in Jus ad Bellum and in Jus in Bello. It is also codified in military manuals and doctrines and, as such, it is widely taught in military academies worldwide. (Colonomos 2017:219)

The use of proportionality as a principle is widespread and yet it remains a contentious principle. Kenneth Watkin opens his paper by saying,

> There may be no other term in international humanitarian law (hereinafter, IHL) which evokes such debate or controversy as 'proportionality'. In part, this debate is a result of the nature of the term itself. The broad use of the term 'proportionality' in international law, combined with images of an almost scientific balancing of opposing interests on finely tuned scales of humanitarian justice, masks a much more complex and unclear reality. (Watkin 2005)

Proportionality is more complex than it might appear, and the term is used in several ways and its usage occupies a broad logical space. The point that Watkin makes is important; although the balancing may be often portrayed as almost scientific, the reality is much messier. We start with a defining characteristic of proportionality: the use of a comparison of the costs of an action to another set of costs related to the action. By clarifying this comparison, and in particular the object of the comparison, we can move towards a technical principle of proportionality that satisfies appropriate normative requirements in the sphere of large-scale conflict. This is the aim of the

following sections, but a small bit of additional terminology is required to help with this aim. Even using the term ‘proportionality’ to describe the entire logical space (rather than a more technical slice of that logical space) may rankle with certain readers, so the term ‘raw proportionality’ is used to describe the logical space and a technical ‘principle of proportionality’ is what we are seeking.

5.2.2. Mapping Raw Proportionality

Raw proportionality, in a normative sense, might be interpreted as a requirement that our actions, given the circumstances, cannot be excessive. Characteristic of proportionality is the comparison of the cost of the circumstances brought about by an action with the cost of another set of circumstances, either actual or counterfactual. What I am terming raw proportionality might be considered a balance between the costs of one scenario with the costs of another scenario. The use of the term ‘cost’ in the previous sentences is not intended to limit the balance to financial issues. The balance of proportionality may contain various types of concern including pragmatic, legal, financial, and ethical concerns. In the arena of moral philosophy, it is common for moral costs to be the dominant factor in the balance, but raw proportionality is able to accommodate a wide range of concerns.

As we may feel that actions are excessive in a range of ways, raw proportionality might be framed in a range of ways. There are three major ways in which one might think an action is excessive.

Firstly, the action may be much greater in scale than the action to which it is a response. This is what Ariel Colonomos (2017:219) describes as a “the tit-for-tat or lex talionis” model of proportionality. It would appear that it is the model of proportionality that Trump was referencing in his tweets that are quoted above. It is a form of retributive proportionality which is often avoided in discussions surrounding war and large-scale conflict. Despite this fact, it is a widely used interpretation of proportionality. In this model, the severity of the response to an action is balanced against the severity of the initiating event. For instance, if I were to gently nudge a fellow traveller in a train, their response is limited to actions that are in some way comparable in severity to a nudge. Raw proportionality would place a limit that would prevent the fellow traveller from stabbing me with a knife in response to a minor nudge. Regardless of whether this form of raw proportionality should be categorised as proportionality, it seems intuitive, and its existence must be acknowledged. For reference later in the chapter I will term it ‘retributive proportionality’.

Secondly, the excess that raw proportionality might seek to avoid is that much more force is used than is necessary. The severity of the action is balanced against the severity of alternate actions that might have comparable results. If a fellow traveller is likely to stab me and I can either prevent that action by shooting him in the foot or in the head, this form of raw proportionality would limit my actions to shooting him in the foot. Again, knowledgeable readers may raise their eyebrows here because this form of raw proportionality is more usually termed ‘necessity,’ or in the case of war studies

‘military necessity.’ However, as is discussed later in this section, it seems likely that the close relationship between proportionality and necessity is best represented as a family membership. At this stage, in which our intuitions or use of language are being investigated, it would be wrong to ignore this form of raw proportionality simply because it has an additional name and some distinct characteristics. It is an intuition against an excessive response based on comparison and such considerations should not be excluded by definition without further justification.

Finally, and as a relief to many readers, raw proportionality can be construed as a balance between the benefits of the action and the harms of the action. This shares a certain amount of ground with a cost/benefit analysis. On one side of the balance is the cost of the action - in other words, the harm that might be inflicted. On the other side of the balance is the benefit of the action - which is often framed in terms of the avoidance of harmful outcomes. McMahan phrases this model of proportionality in the study of war as,

> Proportionality in the resort to war determines a limit to the amount of harm it can be permissible to cause for the sake of achieving a just cause. (McMahan 2015:1)

This definition of the principle of proportionality is used in both legal and international relations contexts. As Israeli ambassador, Ron Dermer said in 2014,

> There you get into the question of proportionality. Meaning, just because it’s a legitimate target doesn’t necessarily give you the right to hit it. Because for that, for you to be able to do that, you have to show that the gain you will get from the military action you take is worth the potential loss of lives that you might even foresee ahead of time. (Miller 2014)

Dermer (and Haque in the earlier quotation) uses the potential loss of lives to represent the expected harm of the action. As this project denies body count morality, I would argue against this part of Dermer’s assessment. Proportionality must not be based on a calculation in which human lives are the only consideration or quanta. However, in fairness to Dermer, it should be noted that the context of his comments was a defence of Israeli actions which were criticised as being disproportionate on the grounds of body count.

These three forms of raw proportionality all provide limitations to the extent of harm that can be inflicted in the pursuit of a legitimate goal. They also use the methodology of balancing one set of circumstances against another set, either actual or counterfactual. In all these cases, what takes its position on one side of the balance is the negative ex-post effects of the action. These are compared either to the negative effects of the ex-ante initiating cause, the negative effects of other methods of achieving the same aims, or the negative effects that are avoided by the action. It is this shared methodology that suggests that all these types of raw proportionality are, at the very least, instances of some common family of comparisons, all of which are used to determine if the negative effects of the action are excessive.

These three examples of comparisons do not exhaust the comparisons that exist in the family. For instance, we might compare against the negative effects of similar historical events which might amount to an 'it was OK last time' type defence which is likely used in practice but has such obvious issues that it is hardly worth rebuffing. The three are the most plausible and widely used forms of raw proportionality. Other comparisons contain such obvious flaws that they can easily be dismissed.

In order to create an appropriate principle of proportionality from what has so far largely been an exercise in investigating the use of language, a process of rational reconstruction[25] is conducted in the following section. By evaluating the normative and political goals of the principle of proportionality, we will be able to narrow the logical space of raw proportionality to a functional and appropriate principle.

5.2.3. Motivating Proportionality

A and B are in conflict about issue Z. The fact that they are 'in conflict' suggests that A and B have conflicting aims with regard to Z. What should we say to A about how they might behave? At one end of a spectrum of advice that might be given is that, in pursuing their aims, A should not inflict any harm at all on B. In this case, the consideration of the harm that may be inflicted is the primary concern and outweighs any other concern. At the other end of that spectrum is the advice that they should pursue their aims irrespective of any harm that might be inflicted on B. In this case, the aims with respect to Z are the primary concern. In both cases, the advice is extreme in that it sits at the end of the spectrum. I will call the first the advice of the lamb, and the second the advice of the lion.

Of course, in an actual or pragmatic consideration of this situation, one is rapidly going to require further information about Z. Perhaps the foremost decision one might want to make with regard to this type of situation is whether one sides with A or B in respect of Z; whether one thinks that A's aims are justified or whether one feels that B's goals are justified. Who has just cause? Or, perhaps, who has made themselves liable to aggression? More likely, one might feel that both goals are somewhat justified, and the decision will be between which is more justified. For the moment, I will attempt to sidestep these concerns in the hope of shining some light on the actual range of possible responses that are delimited by the advice of the lamb and the lion. The aim is not to determine what advice one should give, as this is entirely circumstantial.

[25] Here, I nod towards the Habermasian conception of 'rational reconstruction' but use the term in the manner of Duff (2009:249) when he talks of the use of philosophical technique in the analysis of the law. "This modest conception of the role of philosophy, however, proves to be untenable: clarification must become rational reconstruction — an attempt to make rational sense of the law; and rational reconstruction must involve at least an internal critique, which appraises the law in terms of ends, values or principles that the reconstruction discovers within the law. Such an internal critique must then also point beyond itself, to an external critique that appraises law in terms of the broader and deeper political and moral values by which states should be structured"

Rather, the more modest aim is to determine the characteristics of a principle that allows us to reach the correct evaluation of an action that will cause widespread harm.

The advice of the lamb is unpalatable to many people because it denies the importance of A's aims with respect to Z. The advice of the lion is unpalatable to many people because it denies the importance of the harm that A's actions may create. So, there is a strong motivation to find a middle road between the lion and the lamb. How might we attain the correct balance between consideration of the harm A's actions might inflict and consideration of the value of A's aims?

This mapping of a middle ground is echoed in the structure of Brian Orend's recent book (Orend 2019), *War and Political Philosophy*. He firmly places pacifism, which might be associated with the advice of the lamb, at one end of the continuum and realism, which might be associated with the advice of the lion, at the other end of the continuum. Those options, pacifism and realism, are given a chapter each by Orend (2019:4) as they are "more streamlined and straightforward". However, he continues, "we need more time and effort to explore the complex middle ground which, in a sense, attempts to split the difference," and so, the middle ground gets multiple chapters. The structure of seeking the middle route is characteristic of the study of conflict.

This middle road demands a principle that provides guidance as to how much harm A's actions may inflict given the value of A's aims. In other words, the principle would allow us to know where on the spectrum between lion and lamb the correct balance is between concern for harm inflicted and benefit gained. When the word 'benefit' is used to describe the balance, a similarity is apparent to a cost/benefit analysis in which the cost of a project is balanced against the benefit of the project. In ethical terms, it is clear that what might be required to navigate the middle ground is a moral version of the cost/benefit analysis in which the moral costs of an action are balanced against the moral benefits of that action.

The form of balancing principle that has been widely used is the principle of proportionality. From the very brief analysis above, a criterion of success for the principle of proportionality is that it provides guidance as to how much harm may be inflicted in the pursuit of an aim and it does that by balancing that harm of the course of action against its benefit. In this, a successful principle of proportionality that is used in these situations is forward-looking, in that on one side of the balance is the ex-post harm of an action and on the other side is the ex-post benefit of the action.

It is this forward-looking attribute that separates the required principle of proportionality from the retributive model of raw proportionality. It divides the logical space of raw proportionality and acts as a criterion to select those forms of raw proportionality that are relevant in our principle of proportionality. Returning to the categorisation drawn in 5.2.2, the first model of raw proportionality is excluded while the second and third models of raw proportionality, 'necessity' and 'proportionality' in technical terms, remain as candidates for inclusion in our principle of proportionality.

It is a potentially odd aspect of this forward-looking nature of the principle of proportionality that it is partially blind to the past. That is to say that whether the fellow traveller in the train has previously stabbed you or simply nudged you is not part of the proportionality assessment. What matters is preventing any future injury. The answer to this puzzlement is that this is a technical definition of a principle of proportionality. It does not answer all the difficult moral questions that surround conflict. Rather, it provides one form of guidance as to how those involved in conflict should evaluate their actions. In other words, it provides guidance as to how much harm may be created given the value of a goal. This forward-looking nature explains why proportionality is usually situated as part of a larger framework. It is highly plausible that other factors than this forward-looking proportionality may affect the permissibility of an action. For that reason, in the traditional structure of JWT, proportionality is included alongside just cause and other deontological concerns. Just cause overall has the structure that enables it to engage more fully with past events.

At the start of this subsection, I made the point that I would attempt to sidestep concerns about justice and liability. My more modest aim, I claimed, was to determine the characteristics of a principle that allows us to reach the correct evaluation of an action that will cause widespread harm. The result of this process is that a forward-looking principle of proportionality is appropriate for that evaluation. However, the process is one of simplification and partial abstraction, and it is important to evaluate the results and re-integrate those results into a broader perspective in order to assure that the abstraction has not introduced inaccuracies. As Duff says in the passage footnoted earlier, “Such an internal critique must then also point beyond itself”. The realisation that a technical principle of proportionality must sit alongside other concerns that are able to engage with the significance of past actions is a conclusion of this re-integration.

What we are left with is an understanding that ‘necessity’ and ‘proportionality’ are candidates for inclusion in a principle of proportionality that meets the moral motivations of the principle. The following section provides a more complete analysis of the relationship between necessity and proportionality which results in a more detailed definition of proportionality. This analysis follows the largely theoretical analysis provided by Thomas Hurka. I am not alone in using the foundation laid down by Hurka. His papers have been widely influential in the discussion. In the few cases where I disagree with Hurka’s interpretation, I will note that fact. In any case, it is the structure of his analysis that I borrow rather than the conclusions. Proportionality and necessity form a family of consequence conditions that allow the navigation of the middle ground between the advice of the lion and the advice of the lamb, or, if one prefers, between strict pacifism and strict realism. That middle ground is appealing because neither extreme is entirely palatable to many people in that both extremes show scant respect for interests that may be regarded as significant.

5.2.4. Hurka's Structure of Proportionality

Thomas Hurka starts his paper that is entitled "Proportionality and Necessity" (2008:127) by noting that JWT is not a consequentialist theory "since it does not say a war or act in war is permissible whenever it has the best consequences." Rather, the theory combines deontological and consequential elements. As Hurka notes (2008:127), a theory that did not acknowledge the horrific consequences of war would not be credible as it is "a morally crucial fact about war is that it causes death and destruction." Hurka uses the term 'consequence conditions' to describe the aspects of JWT that concern the consequences of conflict. This seems like a useful categorisation that avoids various interpretations of consequentialism that exist in philosophy. Without needing to engage in consequentialist theories, we are able to see which aspects of JWT are concerned with the consequences of entering into war, or of war itself.

The criteria that Hurka includes in the consequence conditions are Jus ad Bellum proportionality, probability of success, and last resort and Jus in Bello proportionality and necessity. As a complete justification of this group is presented in "Proportionality and Necessity" it is unnecessary to repeat that justification. Hurka (2008:129) subsumes probability of success into proportionality on the basis that in assessing a war that had little chance of success because, inevitably, "its destructiveness is out of proportion to its expected benefits". This is perhaps a little bit quick - the step to expected benefits is undefended and brings with it a raft of complications - but that is not crucial to the following argument.

More crucial is the structure of the analysis of the following criteria. An appealing characteristic of Hurka's formulation is that there is a symmetry in the conditions in Jus ad Bellum and Jus in Bello. Symmetry is always pleasing for its own sake, but in this case, it serves an additional purpose in that it allows that the divide between Jus ad Bellum and Jus in Bello concerns may not be hard-edged. If one is to countenance a softer divide between Jus ad Bellum and Jus in Bello concerns, then it would be problematic to note that our moral criteria were significantly unequal on opposing sides of the divide. The symmetry does not necessarily support the theory that the divide is softer-edged, but it does allow the possibility that that may be true.

To achieve this symmetry, he notes that the Jus ad Bellum principle of last resort equates to the Jus in Bello principle of necessity.

> But in each branch of the theory the proportionality and necessity conditions – the last resort condition is really an ad bellum necessity condition – are independent. A war can be proportionate, because the destruction it will cause is tolerable compared to its benefits, but not a last resort, because the same benefits could be achieved by less destructive means. Or it can be a last resort, because it is the only way of achieving certain goods, but disproportionate, because it will cause excessive harm compared to those goods. (Hurka 2008:129)

So, in both *ad bellum* and *in belllo* we have a proportionality principle and a necessity principle. Further, these two principles are independent of each other yet relate in significant ways. For Hurka, the proportionality principle compares the consequences of an action, in terms of harms and benefits, to the consequences of not acting. The necessity principle compares the consequences of an action to the consequences of alternative actions. Or as Hurka says,

> So in each branch of the theory the proportionality condition considers the relevant benefits and harms of a war or act considered on its own, while the necessity condition compares the result of that calculation with the results of similar calculations for relevant alternatives, allowing a choice only when its balance of benefits to harms is better than that of any alternative. (Hurka 2008:129)

Hurka goes on to claim claims that necessity conditions are derivative of proportionality conditions "because they are comparative versions of them". His point is that in the case of multiple options one would make proportionality assessments for each option and then from this information be able to judge necessity.[26] The important aspect of this discussion moving forward is that necessity and proportionality are members of a broader category of consequence conditions. Seth Lazar makes much the same point when he says,

> Proportionality and necessity are superficially distinct. A war might be necessary, since there is no other means to achieve its end, and yet disproportionate, because the end is not valuable enough to justify the means. To work out proportionality, we need to ask whether the evil inflicted is great enough to justify the evil averted. This means comparing going to war with what would happen if we allowed the threat to eventuate. That comparison is substantively identical to the comparisons between different means for achieving one's ends involved in applying the necessity constraint. Applying the necessity constraint means comparing all your options that have some prospect of averting the threat, to ensure that the chosen one does not involve inflicting unnecessary harm. Applying the proportionality constraint means comparing your best military option with what would happen if you did not avert the threat. Both could be subsumed into a broader criterion, which simply compares all your options in this way. (Lazar 2017:44)

---

[26] Here, I diverge from Hurka somewhat, but in ways that are not critical to the following discussion and may only be linguistic. Rather than believing necessity is derivative of proportionality, I believe that both are derivative of a family of harm/benefit comparisons. There is no hierarchy of derivation between necessity and proportionality. If I am not correct about this, then supporters of Hurka's standpoint owe an explanation of why situations are possible that satisfy either one of the conditions without satisfying the other. In particular, if necessity is derivative of proportionality, it would seem impossible that an action could satisfy necessity and not satisfy proportionality.

In summary, when the term proportionality is used in this project it is taken as a member of the consequence conditions that Hurka describes and may appear either in *ad bellum* or *in bello* contexts. It is a principle that limits the permissible harm of an action by comparison with the benefits of that action, whether that action is a decision to enter a conflict, a decision to continue or escalate a conflict, or a particular action within a conflict. Together with the other consequence conditions, including necessity, it provides the consequential backbone of JWT. In broad terms, it represents the pre-theoretic idea that actions should not be excessive. The problems that will be demonstrated for proportionality apply to all members of the family of consequence condition.

5.3. Espionage

In the Introduction, it was claimed that actions that have direct influence on informational assets are categorised as espionage. This categorisation was based on the influence table. The aim of this section is to investigate the nature of espionage and demonstrate that this categorisation is not problematic or inaccurate.

As before, we can with some dictionary definitions of espionage only to assure ourselves that they do not contradict our technical definition. The OED (OED Online 2022b) is particularly unenlightening, suggesting that espionage is "The practice of playing the spy, or of employing spies." The Merriam-Webster is a bit more forthcoming,

> the practice of spying or using spies to obtain information about the plans and activities especially of a foreign government or a competing company (Merriam-Webster n.d.)

Is espionage defined by what spies do? Let us hope the definition of a spy is not someone who engages in espionage! Here, the OED supplies the more useful definition,

> One who spies upon or watches a person or persons secretly; a secret agent whose business it is to keep a person, place, etc., under close observation; esp. one employed by a government in order to obtain information relating to the military or naval affairs of other countries, or to collect intelligence of any kind. (OED Online 2022d)

It seems that the wide definition of espionage is anything that a spy does, and this is relatively unhelpful by itself. Only when we encounter the phrase “employed by a government to obtain information” do the dictionary definitions become at all detailed. We aim towards the idea that espionage is defined by the unauthorised access to informational assets. The dictionary definitions have not been much help but have not been contradictory to that idea. Perhaps the public understanding of espionage is more helpful?

In conventional terms, the understanding of what espionage is was sealed in the public psyche by the adoption of the image of the spy by popular fiction. The list of fictional spies includes the Scarlet Pimpernel, Smiley and other characters in the novels of John LeCarré, James Bond, Jason Bourne in the films based on the novels by Richard Ludlum, and many others. The icon of the spy is embedded in literature, movies, and most other creative forms but the portrayal of spies in fiction does not necessarily represent the reality of espionage. The famous 'spy' film Three Days of the Condor is a good example.

> This is the prototype for the Bourne films as a CIA agent is ousted by his own people after he learns something he shouldn't and must evade ruthless assassins and calculating spy masters to uncover a large scale conspiracy. (thedryes 2013)

There is an informational aspect to the plot. The initiating event is that the CIA agent learns something he shouldn't. The resolution is informational too, as the discovery of the conspiracy is information. But just those facts would not make a decent movie. We need action and adventure - and those aspects are definitely in the physical domain. In fact, the public perception of the nature of a 'spy' is more related to the media portrayal of characters such as James Bond. This portrayal has very little to do with espionage.

When the Merriam-Webster dictionary says that espionage is "the practice of spying or using spies to obtain information" there is no demand that this is all that the spies do. The spies may be involved in sabotage and subversion (and various other activities) as well as espionage. Perhaps this definition is more useful than it appeared and the critical part of it is that espionage is the practice of using agents to obtain information. This understanding allows us to have a narrow definition of espionage while not objecting to the media portrayal of spies - which focuses on their other activities.

So, the critical definitional element that emerges is that espionage is about gathering information. Intuitively, it is clear that what separates espionage from other forms of information gathering is that it is the gathering of information that the other party would prefer that we did not have. A working definition is that it is a process to gain unauthorised access to information.

In most cases, the process appears to be characterised by covert methods. Traditionally, a spy sneaked into the rival's court and extracted the information in a secretive way. While this might be a characteristic, it is not necessary. Imagine for instance a neighbour builds a platform that allows them to look over the hedge and into your garden. You might object to them spying on you. It would be no defence for them to claim that they could not be accused of spying because their actions were not covert. Likewise, the covert nature of a spy satellite seems very dubious. I am not entirely sure whether these arguments apply to espionage or just to spying. But the fact that there is reasonable doubt suggests that there is not an absolute necessity for covert action in espionage. The reason this is important in the context of cyber

espionage is that many cyber actions are not truly covert - only non-vert or deniable. If there was a demonstrable necessity of covert nature in the definition of espionage it would limit the viability of cyber espionage. I do not think that there is that necessity, so a working definition of espionage becomes:

Espionage is a, frequently covert, process that aims to gain unauthorised access to information.

In terms of this project, an initial outline is provided by the Influence Table that actions that have direct influence on informational assets are categorised as espionage. It is worth noting that this direct influence may involve the unauthorised copying or 'reading' of that information. The analysis of this section has created a working definition of espionage that is that it is a process that aims to gain unauthorised access to information. These two are compatible. They highlight slightly different aspects of espionage but together form a more complete understanding.

Espionage is a process that uses unauthorised access to exert direct influence on informational assets.

It is possible that there are other acceptable interpretations of espionage, and the use of words is a matter of choice. For the purposes of this project, the slightly more technical definition above will be used. It corresponds directly to the Influence Table and is not contradictory with other definitions. What remains to be demonstrated is the nature of that influence and the manner in which cyber methods influence espionage.

5.4. Cyber Espionage

The two aspects of espionage are that it uses unauthorised access and that it uses that access to exert direct influence on informational assets. In this section, those two aspects are discussed in the context of cyber conflict, and in particular using the cyber interests framework developed in Part One.

Database management gives one a valuable framework for thinking about direct influence. Of course, not all information is contained in a database, but they do create a formalised method of storing information. Their ubiquitous nature in computer systems has resulted in a great deal of theorising about the data contained and how it might be acted on. In broad terms, this theory applies to all information, not just that stored in databases and so we can use the computer theory to formalise our thinking about information.

The acronym CRUD is often used to describe the possible actions on data in databases. It stands for create / read / update / delete. In a benign environment, CRUD describes all the things that one might want to do with data. In a less benign environment, CRUD describes how an adversary might wish to exert direct influence on data or other informational assets.

There is some parallel between physical and informational domains in this respect but there are also some differences. In the physical domain, one might destroy an asset which is the equivalent of delete in CRUD. One might alter an asset in some way physically which is the equivalent of update. Creating exists as an action in both the physical and informational domains. However, 'read' is significantly different in the informational domain.

'Read' for a computer scientist implies that the information is copied in some way for use elsewhere in a programme. The original information is not altered or deleted. It is hard to imagine a direct equivalent in the physical domain because the closest possibility is taking a physical asset or removal. Reading in database terms does not involve removal. Perhaps taking a photograph of a document or photocopying it is the closest analogue. Even in that case, there is no doubt that the physical object is not the same as the original even though it contains the same information. It is the information that has been copied (a read action in CRUD terms) while the physical item is not the same. Already, CRUD has clearly highlighted a substantive difference between the physical and informational domains. It is a useful manner to think about information access.

CRUD gives us an understanding of how information may be acted on. Therefore, it gives us an understanding of what might be meant by direct influence. Any element of CRUD may be construed as a direct influence. If an adversary creates new data or information in your systems that is direct influence. Likewise, for updating and deleting data or information. The paradigm example of espionage is reading data or information, but all the other actions are equally types of direct influence.

This cyber interests framework developed in Part One is based on an understanding of the significance of information in our lives, in particular the two interests: private data control and public data access. Cyber espionage can have influence on both these interests.

Private data control demands that there is no unauthorised influence or access to the data over which individuals and groups have ownership. Espionage is in direct opposition to private data control. In paradigm examples of espionage information is accessed and 'stolen'. Here 'stolen' may represent either copying or copying-and-deleting. In either case, the control over the data has been compromised. In less paradigmatic cases, such as creating and updating, which will be discussed in more detail later, that compromise would also be apparent. Private data control is compromised by espionage.

In the case of public data access, it is the less paradigmatic actions that come to the fore. Public data access demands not only that individuals have access to data, but that they have access to data that is reliable. Reliability in this context demands that the information is freely available when required and that it is accurate and appropriately truthful. Creation of new, unreliable information - the 'C' in CRUD - can impact negatively on public data access. This might be demonstrated in misinformation campaigns such as those claiming that vaccines contain microchips

or in the activities of the supposed 'Russian troll farms' that aimed to influence U.S. elections by providing a deluge of inaccurate and skewed information to U.S. voters. Updating existing data - the ‘U’ in CRUD - can have the same effect, but also could be used to undermine common projects such as research and social projects. The limit of how created or updated information might be leveraged is hard to designate and is possibly only limited by the imagination of the adversary.

The ‘C’ and the ‘U’, create and update are traditionally less characteristic of espionage. Our first thoughts about espionage revolve around read or read-and-delete actions. The archetypal example of espionage is the spy who sneaks into the office of the opponent's military leader to take pictures of important documents. This is an example of a read operation. Cyber methods have changed this emphasis because ‘C’ and ‘U’ type actions, create and update, have become much more plausible. This changes the conceptual nature of espionage. It seems unlikely that these opportunities will be neglected moving forward.

Nevertheless, the best examples of cyber espionage at this time remain as read actions. A paradigmatic example of cyber espionage is the extraction of critical military information concerning the F35 fighter aircraft from the U.S. agencies by Chinese infiltrators. That such data was obtained by the Chinese was confirmed by the papers associated with Edward Snowden,

> The Snowden files outline the scope of Chinese F-35 espionage efforts, which focused on acquiring the radar design (the number and types of modules), detailed engine schematics (methods for cooling gases, leading and trailing edge treatments, and aft deck heating contour maps) among other things. The document claims that many terabytes of data specific to the F-35 joint strike fighter program were stolen. (Gady 2015)

This campaign that the Chinese allegedly ran to target U.S. military secrets has been described as wide-ranging and persistent.

> Byzantine Hades is a code name given to a wide ranging and persistent group of network intrusions into U.S. military, government, and corporate systems. The operations can be broken down into three sub-categories: Byzantine Candor, Byzantine Anchor and Byzantine Foothold. Information on the hacks was first disclosed in a report by Reuters, citing information contained in leaked diplomatic cables released by Wikileaks. The cables suggest that Chinese intelligence and military units and affiliated private hacker groups have been penetrating U.S. networks and stealing sensitive proprietary data and otherwise valuable information for years. (Brook 2011)

Byzantine Candor is noteworthy as it led to the widely publicised indictment of 5 members of the Chinese armed forces for the hacks (Sanger 2014). Over a period of time, the Chinese teams managed to exfiltrate a large quantity of actionable information regarding U.S. weapons development. Attorney General Eric Holder acknowledged that states routinely spy on each other but claimed that,

> When a foreign nation uses military or intelligence resources and tools against an American executive or corporation to obtain trade secrets or sensitive business information for the benefit of its state-owned companies, we must say, 'enough is enough.' (Holder 2014)

Effectively, Holder's argument appears to imply that spying against nations is normal and acceptable but spying against a nation's companies is unacceptable. This seems like a strange and fragile conclusion to draw, but presumably stems from the difficulty in ascribing blame for activities in which one is involved. This demonstrates the odd relationship that espionage has with international law. Without an equivalent of the Influence Table, it is hard to categorise espionage, and this has resulted in it becoming, to a certain extent, a grey area for both ethics and international law.

None of the conclusions of the analysis of this type of real-world example contradict our definition of espionage.

Espionage is a process that uses unauthorised access to exert direct influence on informational assets.

The covert characteristic of cyber methods means that they are ideally suited to this type of unauthorised access. It is highly plausible that all the major powers are involved in extensive and long-term cyber espionage campaigns against each other. The recent Solar Winds compromises (SolarWinds 2021) have shown the depth and persistence of the Russian campaign against the U.S. It is unlikely to be an outlier. More likely is that there will be ongoing revelations of such compromises.

In contrast to the previous paragraph, the Snowden papers also demonstrate an attack on private data control. In this case, the action was far from covert. It was inherently overt. This simply demonstrates that 'covert' may be characteristic of espionage but is not necessary.

These examples, as acknowledged earlier, are of read type operations. Are there real-world examples that might support the idea that update and create type actions are available to powers involved in conflict? The answer is that there are few. That may be either because such actions have not occurred, or because they have not been acknowledged. In the case of the F35 fighter aircraft, for example, it might be advantageous for the Chinese to alter the information as well as steal it, in an attempt to degrade the American research efforts. Perhaps, the desire to remain undetected prevents this type of action. Perhaps this type of action is actively occurring but is not acknowledged. No manufacturer wants to cast doubt on the integrity of its design process and systems.

Conceptually the manipulation of data provides such a potential for the projection of influence that its use is unlikely to be ignored. A possible scenario that demonstrates this is the potential manipulation of satellite navigation systems. The American GPS system is used as an example although there are several corresponding systems at this point in time; GLONASS (Russia), NavIC (India), Bei Dou (China) and Galileo

(EU). The development of independent systems has been at least partly driven by fears of loss of control of significant informational resources.

The Gulf Wars were where GPS came of age in the context of wartime operations. Worries about reliance on GPS soon emerged. GPS could, to a certain extent, be jammed. For agents other than those in control of the satellites, there was also the worry that the system could be altered, degraded, or simply turned off. In fact, those in control of the satellites are not immune from those worries because it is clear that those actions may be possible by unauthorised access to the information (e.g., Greenemeier 2016). A 2018 article on the Army Technology (Evans 2018) website highlighted these worries.

> In 2011, Iran claimed to have hacked a US RQ-170 Sentinel drone and spoofed it into landing at an Iranian airbase, while last year, more than 20 vessels near the Russian port of Novorossiysk in the Black Sea found their apparent position had shifted to Gelendzhik Airport – 32km further along the coast. The incident has echoes of the earlier stories of Pokemon Go players around the Kremlin seemingly teleported to Vnukovo Airport, and fuels the growing speculation that Moscow is honing new electronic warfare skills.

If an adversary is able to access one's positioning data and manipulate it, then this constitutes potential degradation to both cyber interests. One's private data control is degraded because the positioning data may be regarded as private to the operators of the positioning system. Public data access is degraded because unimpeded and accurate access to data is protected by this interest.

A question arises as to whether an action in which GPS information is 'spoofed' is truly espionage This question only serves to highlight the potential vagaries between, for instance, espionage and sabotage. Is an action against a GPS system sabotage because it targets the correct function of systems? Or is it espionage because it involves the unauthorised access and manipulation of data? There is no single answer to this question. The two categories overlap significantly. Satellite positioning systems provide a salient example of how information has become increasingly valuable in our lives - and not just for military operations.

Although most of the recorded cases of cyber espionage are of the read type, the potential of create and update operations is large and has not been overlooked. The ability to leverage these 'C' and 'U' operations has been emphasised by the digital revolution and their use is likely to increase.

This section aimed to provide the answer to two questions. What direct influence might consist of in the post-digital-revolution world and how cyber methods and cyber interests affect that influence. The 'CRUD' framework provides a tool for understanding how information might be influenced and has led to an understanding of the changes that are occurring. While read and delete type operations continue, cyber methods allow an increase in create and update methods. The cyber interest framework allows the proper evaluation of the value of cyber interests in this context.

Both private data control and public data access can be damaged by cyber espionage as defined by the definition used in this project. Cyber espionage can create the degradation of interests that are contributory to flourishing human lives and therefore cannot be excluded from the consideration of harm that is used in the principle of proportionality.

Espionage is a process that uses unauthorised access to exert direct influence on informational assets.

5.5. Problems with Commensurability

The previous sections have argued that cyber espionage has significant effects on both private data control and public data access. Meanwhile, proportionality, in its broadest terms, is a concept that aims to prevent actions from being excessive by balancing the harms created against the benefits achieved. For proportionality to be effective and meaningful this balance needs to be possible. The problem arises when the values on one side of the balance cannot be compared with the values on the other side of the balance.

In the cyber interest framework developed in Part One both the harms and benefits are represented by changes in the state of interests. An improvement in interests is taken as a benefit and a degradation in interests is taken as harm. How can we balance the value of private data control with, to borrow Nussbaum's phrase (Nussbaum 2000:41), “Being able to have good health, including reproductive health; being adequately nourished”? However, one phrases it, and nothing relies on Nussbaum's phrasing, there is a problem in comparing interests that are of significantly distinct types. This problem exists, to a certain extent, for all pluralistic interpretations of harm because all types of harm are incommensurable to one extent or another (which partially explains a reliance on body count morality). However, the introduction of cyber interests into the pluralism aggravates the problem to such an extent that the use of proportionality in any meaningful way is impossible.

The fact that the problem exists in all forms of harm is demonstrated by the fact that a traditional way of expressing the Jus in Bello principle of proportionality is to say that the deaths of citizens that are anticipated must be balanced against the military advantage achieved. For instance, although the term ‘proportionate’ is not used, the Geneva Convention prohibits,

> an attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated. (UK Government 2006)

But how exactly should military advantage be balanced against the loss of human life? They appear to be quite different things. The general idea appears to be that certain military advantages are worth a certain number of human lives and, when

framed in that way, this is an unsettling sentiment for many of us. Adil Ahmad Haque describes one of the demands of a successful principle of proportionality in this way,

> … such an account must explain how we can rationally compare civilian losses with military advantages... (Haque 2017:188)

It is worth noting how Haque navigates this question.

> I argue that an attack that inflicts incidental harm on civilians is objectively proportionate only if it prevents opposing forces from inflicting substantially greater harm on attacking forces or civilians in current or future military operations. (Haque 2017:188)

Haque goes on to claim,

> My account of Jus in Bello proportionality is in one way more determinate than existing accounts—for example, it does not compare incommensurable values but instead compares immediate losses to civilians with future losses to civilians and to attacking forces. (Haque 2017:189)

The originality of Haque's standpoint is that the object of the comparison is reduced to "future losses to civilians and to attacking forces" rather than a vague concept of 'military advantage'. His argument is based on the premise that in this way we can sidestep the existing problems of commensurability. The reduction in Haque's method is perilously close to body count morality and only if the 'immediate losses' are of the same type as the 'future losses' is the problem of commensurability entirely avoided. This might be achieved if one thinks only in terms of civilian deaths. If one includes other significant harms or even deaths of combatants, there is no guarantee that these harms are commensurable. Moreover, this problem arises even if one attempts to avoid the use of military advantage in other ways. Commensurability problems exist for all harms. One harm is never exactly equivalent to another. Only if we limit the harms that are to be considered to a very narrow range, in which this lack of equivalence is insignificant, can this incommensurability be neglected.

There is a problem for the concept of proportionality. Fundamentally, we compare harms and benefits. Even if we conceptualise benefits in terms of avoided harms it does not imply that the two sets of harms are commensurable. Only if some homogenising technique such as body count morality, in which the only allowed significant harm is death, is employed does the problem of commensurability vanish entirely. This is a problem that is inherent in the application of the Principle of Proportionality to large-scale conflict but is brought to the fore in the context of cyber conflict.

Two facts explain why cyber conflict emphasises this problem for proportionality. Firstly, cyber interests are significant in the provision of flourishing human lives. Secondly, cyber interests are especially incommensurable with other interests. If it is

hard to compare the value of eyesight to the value of a hand, then it is even harder to compare the value of a hand to an aspect of one's digital body such as private data.

All espionage, but particularly cyber espionage, provides a salient example of the problem because cyber espionage acts directly on informational assets. Espionage is a threat to information assets and informational assets are significant in the provision of flourishing human lives. Informational assets can be individual or can be shared by a group. The result of national research is an example of informational assets that are shared by a group.

In the case of the technical specifications of the F35 aircraft, for example, the nation has rights of ownership over that information and there may be other subgroups such as manufacturers who also have rights over that information. Resources have been committed to acquiring that information and future gain is anticipated from that information. The informational asset has value and a threat to it is a threat against that value.

Imagine that a technologically advanced island state, Country A, relies heavily on shipping to supply its resource requirements. Country B, equally savvy in the cyber realm, conducts a campaign to discover the details of Country A's satellite positioning system - this is a read-type operation. It then uses this information to ensure that the information provided by the satellite positioning system is incorrect - this is an update-, create- or possibly delete-type operation. The first operation, the read-type, is espionage. The second operation, the update-type, is also espionage according to the definition used here because its target is informational assets (position data) rather than systems. (Although this does not exclude it from also being sabotage or subversion.) How can we assess the value to Country A in preventing Country B's plan?

Dismissing the action as 'only about computers' or 'only some information' is inappropriate because the actions are of significance. The standpoint of this project is that the value of information is created by its relationship to human lives. Control of private data and public data access are interests that are not derivative of other interests. This project can define the value in the prevention of Country B's plan in these terms without difficulty. Country B's plan will degrade the cyber interests of Country A and Country A's citizens.

The paradigmatic example of cyber espionage that is used throughout this chapter is the theft of the technical specification of the F35 fighter aircraft. As in the abstract example above, dismissing the action as unimportant is obviously not appropriate. Yet, traditional methods struggle to find ways to incorporate this type of action into proportionality assessments because the target is informational rather than physical. No human lives have been lost and there has been no damage to physical infrastructure. Espionage has always had an uneasy relationship with Just War Theory.

Nonetheless, informational assets have significant value in providing flourishing human lives. So, how else might we find a way of accommodating espionage in proportionality? A pure form of body count morality is of no use here. In fact, the example of espionage demonstrates the weaknesses of this type of approach.

Perhaps we might rely on an abstraction such as ‘military advantage’? In the case of the F35, it might be possible to say that there was a military advantage associated with the technical information. This seems plausible. However, the problem in cyber conflict is that the informational target may not be associated with the military. The satellite positioning system used in the abstract example above may have military uses but there are other non-military uses such as transport that may have greater significance. The target of the action is not necessarily military. Nor are the actors necessarily military. In the previous chapter, we saw that cyber conflict was not the exclusive domain of the military. So, the concept of ‘military advantage’ is inappropriate in most cases.

How about using 'advantage' rather than 'military advantage'? The problem that we are trying to overcome is that various harms are incommensurable. All types of action might be said to have advantage - and so perhaps advantage is a concept that might be used to assess incommensurable interests. Perhaps 'advantage' can homogenise incommensurable harms? It might be possible that the value of the informational asset can be assessed by the advantage it provides in the same way as the value of a physical asset might be evaluated by the advantage it provides. This idea is put aside for a few paragraphs.

First, is there another way in which the side of the balance that is traditionally represented by 'military advantage' can be expressed? Haque’s idea was that one might use 'future losses to civilians and to attacking forces' as a metric. He anticipated this being used as a replacement for 'military advantage' on that side of the balance. How does that cope with the examples?

Possibly one might be able to claim that the theft of the technical specifications for the F35 would lead to future losses. The air superiority that the F35 might otherwise provide has been diminished. The argument would go something like this: the Chinese would be able to carry out actions of which they otherwise would not be capable, and at least some of these actions would be disproportionate - so on that side of the balance we would be able to place the losses that those actions caused. In traditional terms, we might say that the Chinese actions would be anticipated to cause the loss of 100 civilian lives, or some such estimation.

The problems here are myriad. The first is that the Chinese will claim that (even if they did undertake the espionage) none of their anticipated actions are going to be disproportionate or unjustified in any way. Perhaps we can just ignore that claim on the basis that all military action tends to have collateral damage? That is not a favourable route for a military country because it results in severe doubt about any military action. The second problem is that the anticipation of 100 civilian deaths is largely arbitrary. More likely is a belief that some civilian deaths will occur. This only

adds to the vagueness of 'future losses to civilians'. The third problem is that we still need to define what type of 'future losses' are going to be included in our assessment. Do we include deaths? Do we include injuries? Do we include property damage? And, finally, do we include informational damage? This third problem is the final conceptual block to Haque's method. The use of anticipated future damages does not resolve any problems as to how we might define losses. It still requires an abstraction of harm - such as body count morality - to be effective. Even if we use some sort of threshold approach in which only serious harms are considered, we are left with the problem of defining what a serious harm is. None of the intrinsic problems of proportionality are resolved; they are simply displaced.

Haque's approach is a serious attempt to confront the intrinsic problems of proportionality. Rather than confronting those problems, the approach moves the tension to another area. The question of how to define 'civilian losses' is a parallel to the problem of how to define harm in general. If one accepts the need for a pluralistic appreciation of harm, it is inevitable that one will require a pluralistic appreciation of civilian losses. Any pluralistic appreciation of harm will contain potentially incommensurable interests. Of course, it is tempting to resort to a technique based on abstraction of harm, but this type of approach has already been discounted in this project.

The initial question was how we might accommodate espionage in proportionality. This was an example of the problems posed by incommensurable interests for proportionality. It is an example that espionage demonstrates because espionage relates directly to informational assets. In the analysis above, the only promising avenue was to assess the significance of espionage - the weight it puts on the balance - by using the concept of 'advantage'.

'Advantage' is a somewhat vague concept. Perhaps, advantage would represent the good that could be achieved with possession of the information asset? This certainly seems relevant. If I am involved in an 18th century conflict and know how to make gunpowder that is certainly an advantage. So, information creates advantage in this case. The Oxford English Dictionary (OED Online 2022a) defines advantage as "Benefit; increased well-being or convenience."

How then does the use of advantage help avoid the problems of incommensurability? The hope is that advantage is a concept that is common for even incommensurable interests. Rather than evaluating the value of the informational asset itself, one evaluates the benefits it might produce, and one hopes that these benefits are not incommensurable with the benefits of other assets or actions. To reinforce this point if advantage is increased well-being or convenience then disadvantage is decreased well-being. Decreased well-being is an analogue of harm.

If disadvantage and harm are analogous, then using advantage to represent a side of the balance in proportionality is analogous to using harm. It is impossible that this avoids any of the problems associated with harm. And in no way allows the comparison of incommensurable interests. Their benefits, advantages, harms, and

disadvantages may all remain incommensurable. The use of advantage as a metric does not avoid the problems of incommensurability.

The use of 'advantage' was our best hope at answering the question that was posed earlier; how else might we find a way of accommodating espionage in proportionality?

This question arises when a plurality of significant human interests is accepted. Those interests are likely to be incommensurable to a greater or lesser extent. Where those interests are incommensurable there is no other technique or derivative value that can be used to assess their relative worth.

While this problem exists for proportionality in general, it is brought to the fore with regard to cyber interests. The cyber interest framework correctly assigns value to cyber interests. It is the case that informational assets and interests are likely to be incommensurable to a greater extent with physical interests than physical interest with physical interest. Espionage provides a clear example of this. The advantage, or disadvantage, of espionage is clear. Its effects can be wide-ranging and can alter the trajectories of conflict. Yet, the target of espionage is informational assets that are largely incommensurable with physical interests.

Proportionality cannot be correctly applied in situations where the interests involved are incommensurable. At the most fundamental level, proportionality is a comparison of relative value. If the relative value cannot be assessed, then proportionality is of no value or use.

5.6. Rescuing Proportionality

The conclusion of the previous section, that it seems unlikely that proportionality can be applied in a principled fashion, is uncomfortable to many. Do we have to bite the bullet, or are there plausible routes around this problem for proportionality?

We might accept that incommensurability poses a problem for proportionality but believe that it is not a terminal problem. This is the route that appears to have been adopted by most theorists historically because proportionality remains positioned as a core element of theories of conflict. There are two ways in which this might be justified. Firstly, it might be claimed that although harms are technically incommensurable, human judgement has the necessary abilities to overcome this problem. I call this the judgement approach. Secondly, it might be claimed that although the harms are incommensurable, the overall harm tracks one or other of the harms. This is a parallel to one of the defences of body count morality that was discussed in Chapter Two, and here I will call it the tracking approach.

The judgement approach starts with the thought that although harms may be technically incommensurable, we believe that we are still able to make an acceptable assessment of them by using human judgement.

There are some grounds for believing in this reliance on human judgement. The idea that we can balance one type of value against another exists in a wide range of situations. In deciding which of two sub-optimal paths we choose to follow we often balance harms and benefits against each other. In cases where the harms or benefits are incommensurable, we often rely on 'feel'. We do not know, for instance, whether it is more beneficial to live in a nicer house further from work or a less nice house closer to work. Proximity to work and comfort of the house are largely incommensurable, and yet we still manage to make these decisions.

This intuitive balancing of harms/benefits appears in various contexts including international reparations, retributive punishment, and industrial compensation. For example, in the case of industrial injury, there are well-defined scales of compensation for loss of a limb. Currently, the loss of a finger is usually compensated by a payment of around £15,000. So, although the benefit of having the money is technically incommensurable with the loss of the finger, we have a rule of thumb regarding their relative value (in passing, the loss of a thumb is equivalent to about £25,000). Here, I am using 'rule of thumb' to signify this kind of formal understanding or relationship, but also a less formal pre-theoretic belief that one harm is roughly equivalent to another. Overall, it is this kind of intuitive judgement that props up the Principle of Proportionality. We must ignore the fact that almost all harms are ultimately incommensurable and concentrate on our supposed ability to make intuitive assessments as to the relative value of certain harms and benefits. Whether or not this form of rescue is valid in the context of cyber conflict is discussed in the following section.

According to the judgement approach, intuitive assessment of values appears at the heart of the Principle of Proportionality, a principle which in turn is at the heart of much legal and moral theorising concerning large-scale conflict. To a certain extent, this may explain the distrust some theorists have in the supposed clarity of Just War Theory when applied to the less certain and more malleable environment of the real world. The Kenneth Watkin quote that was used earlier is an example of this.

> The broad use of the term 'proportionality' in international law, combined with images of an almost scientific balancing of opposing interests on finely tuned scales of humanitarian justice, masks a much more complex and unclear reality. (Watkin 2005)

As Michael Neu says of the philosophical analysis of war,

> it often deals with neat fictitious worlds, rather than the complex real world; it makes clinical moral judgements about large-scale killing, mutilating and suffering, and it proceeds as if absolute moral and epistemic certainty could always be obtained. Contemporary just war theorists have turned the world of war into a sterile analytic playground. (Neu 2013:462)

Cheney Ryan makes much the same point when he says, talking of contemporary JWT,

> But its vice is its abstraction. This is especially worrisome given that ideologies of war have often appealed to war as imagined, rather than war in reality. Reading just war theory, it is easy to feel that one is encountering a lengthy, complex discussion of what "breaking and entering" means—that never stops to ask why every house in the neighborhood is being vandalized every night. (Ryan 2013:979)

Nevertheless, the judgement approach needs proper consideration because it is a possible route by which proportionality might be defended. It is the subject of the following section. The tracking approach on the other hand can be quickly dealt with. It can be used as a stand-alone defence of proportionality, or as an extra prop for the judgement approach. Either way, it is worth analysing what it says.

The tracking approach claims that even if some set of harms/benefits are incommensurable with another set, the one set generally tracks the other. So, in the case of body count morality it might be claimed that although severe maiming is significant and should be included in our proportionality assessments, it can be ignored because it correlates to physical death. That is to say, in a conflict where we have X number of deaths on one side and Y numbers of deaths on the other, the number of maimings on the first side will be in approximately the same ratio (X/Y) to the number of maimings on the second side. We can therefore ignore maimings because they track deaths. This logic might be extended to encompass harms related to all interests, including cyber interests.

In order to disprove the tracking approach, one would need to find examples in which the ratio of one type of significant harm did not correlate with the ratio of another type of harm. This correlation between other harms and, for example, the number of deaths is challenged most strongly by widespread dispersed harms in large-scale conflict. Imagine two scenarios. In the first, coercion is attempted by enemy forces and involves the focused application of intense physical harm including death. This is the conventional model of war. In the second, coercion is attempted by the widespread application of dispersed harm over a population. In this model of conflict there are significantly fewer deaths but many more instances of lesser harm to individuals. As was argued in the second chapter this does not imply that there is less overall harm involved in the second model. In these circumstances, it is clear that the other harms do not track deaths. Again, as has been argued previously, cyber conflict is by character, persistent, dispersed, transversal and covert. These characteristics imply that much of cyber conflict will be closer to the second model of conflict than the first and, therefore, it is unlikely that other harms will track the number of deaths.

It is likely that in the context of cyber conflict a tracking approach will become increasingly inappropriate. As shown in previous subsections, there is an increasing distinction between the informational and the biological body. In particular, cyber methods allow influence and coercion to be applied to informational aspects of lives without necessarily applying influence to biological lives. This is a substantive and significant difference because, without the close link between the two aspects, any tracking approach is unlikely to succeed. Moreover, the nature of cyber conflict has

been shown to be dispersed rather than tightly focused - more like the second model of conflict in the paragraph above - and this too makes belief in the relevance of the tracking approach difficult.

At this point, we can dismiss the significance of a tracking approach. The very nature of incommensurable harms suggests that it is deeply flawed. We are left with the possibility that the judgement approach might be appropriate if we can find a homogenising metric such as 'advantage' that might overcome the problems of incommensurability. Considering that idea is the subject of the next section.

5.7. The Judgement Approach

The problem for proportionality is that almost all harms of different types are incommensurable. Even within the biological aspect of our lives, harms are difficult to quantify and compare. Is losing a hand equivalent to losing a foot? How many fingers lost would be the equivalent of losing one's sight? Do we need air more than food? For our purposes here, the claim that I am making is that across different aspects of our lives, comparison is impossible. So, as the aspects that are core to the argument here are the biological and the informational, the claim is that it is impossible to compare biological factors to informational factors. The question 'how many fingers equates to one's informational history' is unanswerable in any coherent and principled manner.

If the points of the previous paragraph are taken seriously, the sceptic of my argument will ask how it is that proportionality has been of any use historically, or why it is that proportionality has formed such a key part of our legal and moral frameworks. They will claim that while we cannot assess the comparative values of certain harms in a scientific manner, we are more than capable of making reasonable assessments of the relative value of harms. 'If I ask you whether I should cut off your earlobe or your leg,' they might suggest, 'it is clear that people will opt for the earlobe.' This kind of assessment, they claim, is sufficient to rescue proportionality. The claim that the sceptic must make is that human judgement is sufficient to compare these incommensurable harms in an acceptable fashion.

I term this the judgement approach to the defence of proportionality. It is a reliance on intuition to supply the correct answer to a proportionality assessment. That intuition may be pre-theoretic, such as the assessment that a leg is more important than an earlobe. It may also be more formalised, as in the case of legal compensations described previously. Whether formal or informal, a judgement approach is based on a process in which the relative merits of outcomes are compared using intuition, human judgement, or other subjective techniques.

To apply proportionality in large-scale conflict we need to be able to assess harms of great magnitudes and over large populations. Our intuitive judgement must be functional at these scopes. There is a legitimate worry that at these scales our judgement breaks down.

There are two possible routes for the defence of the judgement approach in large-scale conflict. The first is that we can intuitively comprehend the magnitude of harms of relative actions in war and make appropriate ethical assessments directly. A more common route is to investigate our ability to apply the judgement approach at smaller scales and then argue that these conclusions can be scaled to war or other large-scale conflict. This second route is the one that will be tackled in the following paragraphs. It is fortunate that the arguments against this route would also provide ample evidence against the first route.

There are two steps to this route. The first is to demonstrate that the judgement approach works at smaller scales. The second is to demonstrate that this conclusion can be scaled to accommodate the largest of scopes. Both are discussed and challenged below.

5.7.1. Step 1: Judgement at Smaller Scales

There are two conflicting sentiments that are appropriate. The first is that in everyday life we feel that we are capable of making assessments of the value of certain outcomes on an individual or small scale. The second is that we are more than aware of the fact that our judgement can be flawed. Which of these voices should we listen to?

If we were unable to weigh up contrasting options in terms of harms, then we would never be able to make a decision. The fact that we do make decisions suggests that we have a method of assessing even incommensurable values. However, the fact that we do occasionally regret decisions that we have made suggests fallibility in these skills. In the example used above, in which a decision between the comfort of a house and its proximity to work, it is true that most of us would be able to make a decision by deciding which felt right. That is not to say that we might not regret that decision subsequently. In fact, any difficulty that we might experience in making the decision suggests that there is the potential for error. The truth is that we are very prepared to acknowledge that we might get it wrong even if ‘it felt right at the time’. This truth speaks against the claim that we have a power of intuition that allows us to overcome the problems of incommensurability. Perhaps then, our power of intuition can be exercised carelessly and needs to be exercised correctly? However, if this were the case, it seems likely that we would never find ourselves in a position of regretting a decision. We would simply need to ensure that we exercised our intuition correctly. It seems unlikely that we have an infallible skill at the smallest of scales to use a judgement technique successfully.

When we turn the spotlight on other ability to exercise this supposed power of intuition, our judgement is increasingly harsh. We might believe that we, personally, get it right most of the time, but it seems less likely that we can say that unreservedly about others. And this is the case even if we share the majority of our values with those other people. If we do not share the majority of our values, then the possibility

of believing that others will always make correct proportionality assessments recedes even further as is discussed in subsection 5.7.3.

Perhaps, even if we do not have an infallible skill at assessing the balance of benefits and harms, our skill is ‘good enough’? It would seem that the final judgement on whether it is good enough would depend on one’s risk tolerance and the extent of the risk. In the example used previously, buying a house, the jeopardy is relatively low. It makes an example of an assessment involving incommensurable values but is not an appropriate example if we are going to scale up our conclusions to large-scale conflict. An example with even slightly greater jeopardy is required. Imagine, for instance, that we can save a neighbour’s highly loved cat - but only at the cost of losing a finger. Do we have the skills to make a judgement type assessment in such a circumstance? As the jeopardy increases, we might feel that our intuitive abilities become less secure.

A conceivable response at this point is to argue that there are scenarios in which there is no one correct answer. That the correct answer is indeterminate (e.g., Elster 1989). Even if this is the case, which is very plausible, it only reinforces the point that personal judgement is not always able to provide us with a correct assessment.

It is sufficient at this point to argue that there is some doubt as to whether our ability to use a judgement approach is practical at small scales. A sense of doubt is all that is required at this point. The advocate of this type of intuitive judgement in large-scale conflict is going to have to build on this foundation and scale up the conclusion to situations with the greatest jeopardy. The real challenge to the judgement approach will come with the process of scaling up itself, but it is worth noting that the foundations of this scaling up are themselves dubious because it is not clear that our human judgement is sufficient to make appropriate judgements even at the personal scale

5.7.2. Step 2: Scaling Up

The problem of scaling up ethical judgements is the subject of a complex discussion that underpins the division between individualist and communitarian approaches to the theory of large-scale conflict discussed in the first chapter. It is not the aim here to enter into that discussion. The aim is to discuss two major challenges that scaling-up has in the precise context of large-scale conflict. Those challenges are an understanding of aggregation and a difficulty in understanding the magnitudes encountered.

Even if we pretend that we have an adequate ability to make intuitive judgements between incommensurable values at a personal scale, the challenge is to apply these judgements at a larger scale. The hope is to extrapolate from this individual scope an understanding that is appropriate to large-scale conflict. I call this process 'scaling-up' and it is more complex than is sometimes assumed.

Regarding cyber interests, the scaling up problem can be seen if we look at the value of private data control. It has been argued that private data control is an interest that is significant in a flourishing human life. From this, it is clear that private data control

for an individual has a certain value. One might think that ethical assessments based on this value are possible. For instance, an outcome with a greater control of private data might be regarded as preferable to other outcomes. Even on this individual scope proportionality is challenged. Is a particular increase in private data control proportional to a specific harm such as the loss of a finger? Even if we do have some idea of the answer to that question, can this understanding of the value of private data control simply be transplanted to groups larger than the individual?

Two factors stand in the way of confidence in scaling-up. The first is that to scale-up in a principled way, we would have to be assured that the process of scaling-up was coherent. Challenges to this cluster around how the value of a harm repeated multiple times relates to the value of that harm on the individual level - in other words how do we aggregate harms? The second is that we would have to have a clear understanding of the scope to which we were scaling-up. Without this, there is no way to determine if the first factor is satisfied. The main challenge to this is the fact that we are poor at comprehending large numbers of the sort that are involved in large-scale conflict.

Aggregation

In order to have confidence that the scaling-up process is coherent one needs either to have a principled way to aggregate harms or benefits or, at the very least, be very clear as to the form that aggregation might take. This is because scaling-up inevitably is grounded by aggregation of harms which implies that the manner of the aggregation is important.

Taking the example of private data control used above, it was claimed that we understand that a change in private data control for a million people is more significant than the same change in an individual. Is it possible to deny that understanding? It is not plausible to do so in any coherent way that satisfies the basic equality of humans. If we accept that it is more significant, then we are aggregating the harms in some manner.

There really is no consensus as to how harms should be aggregated. It might first appear that aggregation of harms is simple. We should simply add the harms together to provide an overall idea of the total harm. This might be regarded as pure aggregation. There are strong arguments, both in philosophy and in public belief, to doubt pure aggregation. The majority of these arguments are based on the idea that no amount of lesser harm may equate to or balance a much greater harm. For instance, the avoidance of a large number of papercuts can never justify a single death, regardless of how large a number is involved.

Much of the discussion surrounding aggregation has concerned hybrid or intermediary types of aggregation. This has been driven by the thought that aggregation is intuitive when the harms are relatively similar, but not when the harms are very different. Pat Tomlin (Tomlin 2017:232) provides useful examples of this type of argument:

Tomlin 1.
You can save one person from death or some larger number of people, N1, from paralysis.

Tomlin 2.
You can save one person from death or some larger number of people, N2, from a mild headache.

The idea here is that many people believe that there is a realistic value for N1, the number of cases of paralysis, that balances or outweighs the single death but also believe that there is "no number of mild headaches that could outweigh, in the relevant sense, the importance of saving someone's life" (Tomlin 2017:233). Those, who I will term the advocates of limited aggregation, who support these intuitions are burdened with explaining why some harms 'matter' and others do not.

One of the clearest explanations of the issues surrounding aggregation is by Paul Voorhoeve (Voorhoeve 2014 & 2018) - who supports a form of limited aggregation. Voorhoeve presents a summary of relevant empirical research into these issues. The issue that Voorhoeve picks out is that a significant number of people demonstrate the combination of intuitions that might motivate a limited form of aggregation - that in situations of relatively similar harm they support aggregation, while in situations of disparate harm they deny aggregation. However, what is seemingly glossed over is that the majority of people do not have this combination of intuitions. There are a significant number of people who believe that an aggregation of small harms can balance a number of deaths, but they are not the majority.[27]

The purpose of this overview is not to resolve the issues surrounding aggregation, but rather to cast doubt on any idea that aggregation is without its concerns. Overall, the aim of this section is to respond to attempts to rescue proportionality by the rule-of-thumb approach. The outcome here is that there is some doubt about how harms can be aggregated.

One answer to this doubt might be to claim that the search for a principled form of aggregation was misguided. The line of that argument goes something like this. Scaling-up does not have to be based on a principled process of aggregation. That is the point. We are not aggregating. We are saying that we know that one headache is better than one death. The lesson we learn is that in cases of equal numbers of headaches and deaths, it is better to choose the option with headaches.

This rule-based approach to scaling-up makes perfect sense - up to a point. We might simply be able to use those rules in all cases, including cases of a million deaths and a million headaches. But those cases are not problematic in any way. The problematic

[27] For instance, Voorhoeve says (2018:128) "These findings suggest that a substantial minority (35.7%) believed both: (a) that no number of cured cysts can outweigh saving several lives; and (b) that curing a large number of meningioma cases should take priority over saving lives"

cases are those in which a million headaches are opposed to a single death. The rule-based approach fails here because the smaller scale example is balancing one headache against a millionth of a death, which is an incoherent concept.

Aggregation poses extreme challenges for the scaling-up process. The advocate of scaling-up owes us an explanation of how the process works and why it is that conclusions on the small scale are necessarily appropriate on the large scale.

Large Numbers

As described above, a second challenge to the process of scaling-up is that we need a clear understanding of the level to which we are scaling-up. The previous subsection used Pat Tomlin's example that there is "no number of mild headaches that could outweigh, in the relevant sense, the importance of saving someone's life". This subsection will argue that a possible explanation for this supposedly common intuition would be that human beings really have no adequate comprehension of large numbers - we just don't get the concept of a million headaches because we just don't get a million. So, we do not have a clear understanding of the level to which we are scaling-up.

The aim of this section overall is to challenge the application of a rule-of-thumb approach to proportionality assessments regarding large-scale conflict. Earlier, I stated that there were two routes for the defence of the rule-of-thumb approach, the scaling-up route, and the direct assessment route. If it can be shown that human judgement does not cope well with large numbers, then this is a roadblock in both routes. For the direct route - in which we make ethical judgements based on the actual circumstances of large-scale conflict - a comprehension of the numbers involved would be essential. Even for the scaling-up route - in which conclusions from a smaller scale are imported into consideration of large-scale conflict - and understanding of large numbers is also necessary. Without an understanding of these sort of numbers the process of scaling-up cannot be justified.

I have argued that our intuition is not sufficiently powerful to overcome the problems of incommensurability of harm on the individual level. In the context of large-scale harm, the harm may well be inflicted across large groups, possibly millions of individuals. In order to be able to start to assess this type of harm, we would need to be able to properly comprehend the large numbers involved. However, there is a problem here, because as Nobel Laureate, Daniel Kahneman says,

> Human beings cannot comprehend very large or very small numbers. It would be useful for us to acknowledge that fact. (Adams 2012)

Steven Pinker describes Kahneman's overall contribution as follows,

> His central message could not be more important, namely, that human reason left to its own devices is apt to engage in a number of fallacies and systematic errors, so if we want to make better decisions in our personal lives and as a

> society, we ought to be aware of these biases and seek workarounds. That's a powerful and important discovery. (Pinker 2014)

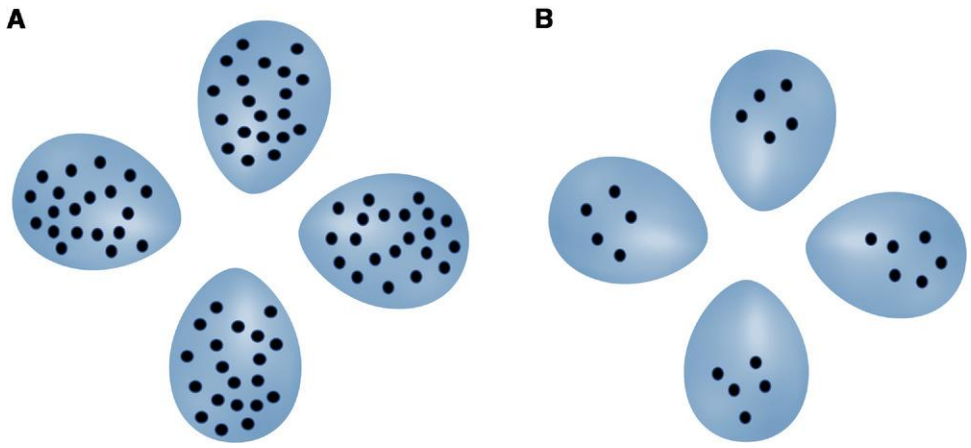
In this section, the aim is to map out one of those failings in human reason. That particular failing might be described as a systematic error in the comprehension of large numbers. As will be seen, this is a specific failing that is discrete from other potential failings. Its particular nature is that we do not have a reliable method of comprehending the magnitude of large numbers in relation to each other. Instead, we rely on the language representation of the numbering system to comprehend large numbers which does not necessarily equate to their magnitude. We know that one plus one is two and can claim to fully comprehend the magnitude of two. This does not mean that we necessarily can comprehend the magnitude of a million, despite the intellectual knowledge that it is a thousand thousand.

'Comprehend' as used here does not equate to having the vocabulary or mathematical skills to interact with the concept of large numbers. Rather, it refers to an ability to intuitively understand or ‘get’ the magnitude of large numbers. As we will see below, one way that this type of comprehension has been described is our ability to correctly place numbers on a number line.

> Human adults are thought to possess two dissociable systems to represent numbers: an approximate quantity system akin to a mental number line, and a verbal system capable of representing numbers exactly. (Izard and Dehaene 2008:1221)

We might know how one thousand relates to one million, and how one million relates to one billion, but when asked to place large numbers on a number line the results are usually highly inaccurate.

At this point in the conversation, Weber’s Law is often invoked as an explanation. The Weber-Fechner Laws relate the ratio of the perceived change in a stimulus to the actual change in the stimulus. Our perception of change is proportional to the absolute magnitude of the input. So, for example, in each of the groups of eggs in the diagram below (Dixit, Tanmay, et al. 2021:1907), one of the eggs has more spots than the other eggs. In the right-hand group where the absolute magnitude is lower, the exception can easily be recognised, whereas with a greater overall number of spots the exception is harder to recognise.



Weber's Law does not reference the type of 'comprehend' that is being used in this section. It refers to our ability to discern the difference between two groups visually. This is different, although possibly related in important ways, to our ability to comprehend large numbers.

In the quotation above, a contrast was drawn between a mental number line and a verbal system. While we might understand that a million was a thousand thousand because we have a verbal system, that does not necessarily imply that we comprehend a million or that our mental number line is an accurate representation of large numbers. Landy et al. put the same point somewhat differently when they claim that,

> understanding the magnitude of 2 billion might be less like perceiving its quantity than like believing a system of facts that involve magnitude systems (Landy et al. 2014:819)

Their research demonstrates that any mental number line that we have is flawed in that it is likely discontinuous. That is to say that our comprehension of how numbers are related to each other is not correct and has discontinuities - particularly around culturally important points such as one million. They conclude that,

> A fundamental mistake made by classical empiricism was to assume that the inner representations were iconic—that they were like the outer represented. When reasoning about large numbers, we appear to rely on representations that are fundamentally unlike the numbers themselves. (Landy et al. 2014:820)

In the simplest terms, the research shows that we do not have an accurate grasp on the gap between, for example, one thousand and a million in comparison to the gap between a million and a billion. We believe the gaps are more similar than they actually are.

The paper by Landy, David et al. specifically targets the comprehension of ‘big numbers’ around a million or more of which we have little direct life experience. The suggestion (Landy et al. 2014:815) is that "There is little reason to think that evolution would have specially prepared us to deal with quantities of this magnitude." Numbers of this scale do not occur in our everyday lives, but they may be important for studies such as "geology, astronomy, and macroeconomics, and in political contexts such as budget discussions." Most importantly for this project, they are the scale of numbers that are critical in ethical discussions regarding harm in large-scale conflict.

This research corresponds to a widely held belief that "We humans are a smart bunch, but we really suck when it comes to understanding and handling excessively large numbers."[28] An online search will quickly reveal support for this concept. In this respect, the conclusions of the research on our mental number line are not controversial. Both the research and the popular zeitgeist show that we are not good at comprehending large numbers.

The initial example that was used of an extreme proportionality assessment was that there might be “no number of mild headaches that could outweigh, in the relevant sense, the importance of saving someone’s life”. Let us say that the number of headaches is a million. The problem is that we do not have a clear grasp on what is meant by a million headaches.

A rule-of-thumb approach accepts that proportionality is not a 'scientific balancing' but instead requires that the relative merits of outcomes are compared using intuition, human judgement, or other subjective techniques. One cannot assess the merit of an outcome involving a large number of individuals if one does not have a true grasp on the magnitude of that large number. Even if our general moral intuitions are not flawed, the results of any assessment cannot be trusted if the magnitude on one side of the balance is not adequately comprehended.

Can we scale up successfully?

Proportionality assessments in the context of large-scale conflict have significant import. In this type of conflict, the potential harms are extensive. If we are going to base our trust in proportionality on a process of scaling-up then that process must be demonstrably reliable. This is not the case.

Firstly, from the previous subsection, we know that our judgement at smaller scales, which forms the base for scaling-up, is not solid. Secondly, the process itself inevitably involves some form of aggregation and there is no consensus as to how this should be carried out and whether it can be effective. Thirdly, an inability to comprehend

[28] The website quoted (Dvorsky 2014) is representative but is neither endorsed nor denied. It is simply used as an example of this type of information that assumes a difficulty in comprehension.

large numbers causes problems because we don't really understand how far we are trying to scale-up.

None of these three challenges provides a terminal body-blow to the idea of scaling-up ethical judgements in this context. However, the onus is on the advocate of scaling-up to explain how they can be overcome.

If we have reduced confidence in our ability to scale-up, we may claim that we are able to intuitively make assessments at the large-scale without basing them on smaller scale assessments. This direct judgement suffers the same two challenges as a scaling-up process with regard to large number comprehension in particular. It seems implausible to accept that we do not really comprehend large numbers but claim that we can correctly assess the negative impact of harm over a large number of people.

5.7.3. Reliance on Universalism

The main thrust of this section is encapsulated in the two previous subsections. Our ability to judge at smaller scales is less than perfect and that ability cannot be scaled up to larger scales. Although it is not the focus of this section, it is worth noting a further challenge. The idea that judgement might be used in proportionality assessments regarding large-scale conflict assumes that there might be a single correct judgement at this scale, or at least a single widely acceptable judgement. This implies a level of universalism of values that is questionable.

It is worth rereading the quotation from Kenneth Watkin,

> The broad use of the term 'proportionality' in international law, combined with images of an almost scientific balancing of opposing interests on finely tuned scales of humanitarian justice, masks a much more complex and unclear reality. (Watkin 2005: extract)

If this form of scientific balancing were possible then proportionality would be safe from the criticisms that I am levelling in the previous subsections. However, it is not, and a judgement defence requires that the relative merits of outcomes be compared using intuition or human judgement. A judgement approach is inherently subjective rather than scientific – that is, it depends on a personal set of values. There is no guarantee that there will be any adequate universality in those values in the relevant areas.

Different people will have different value profiles. People on opposite sides of large-scale conflict will inevitably have different value profiles. The difference may be determined by political beliefs, religious beliefs, or by other factors such as wealth, education, or geographic location. Different value profiles will result in different proportionality assessments.

The defender of proportionality may claim that there is adequate universality in the most important factors of proportionality assessments. The onus would be on them to demonstrate that this is the case. Certainly, there are strong arguments against this type of universalism in various literatures.[29] This is one area in which this project has the luxury of not engaging. The conclusions of the previous two subsections stand independently of this issue. Their purpose was to conclude that the judgement approach could not be justified, and the challenges of universalism can only strengthen this conclusion.

5.7.4. Subsection Conclusion

For a judgement approach to proportionality to be successful one must be able to hold that human judgement is sufficient to compare harms in an acceptable fashion even if these harms are technically incommensurable. Moreover, this sufficiency must be displayed on the greatest scale, in which perhaps millions of lives are adversely influenced, and not just on more parochial and personal scales.

It is not clear that we as humans have infallible judgement of incommensurable interests even at the smallest of scales. Particularly, small increases in jeopardy can lead to doubt about this judgement. Even if we trust our own judgement, we often trust the judgement of others less.

Even if our judgement was adequate at lesser scales, it is not clear that the ethical assessments that are made at this scale can be extrapolated to larger scales. Problems of aggregation emerge quickly and remain unresolved. Problems of comprehension of the scale of the target of extrapolation due to poor comprehension of large numbers mean that it is impossible to assess the success of the extrapolation.

The judgement approach is an inherently subjective approach. There is no evidence that the relevant beliefs are universal enough to make the judgement approach widely acceptable. It seems more likely that the judgement approach can be leveraged to

---

[29] For an overview of universalism in ethics, see O'Neill 2016 && Rasmussen 1990. For an example that is pertinent to the previous chapters refer Martha Nussbaum's defence of her interpretation of the capabilities approach against charges of ethnocentricity. Nussbaum (1995:5) claims, "The capability view is in principle compatible with cultural relativism—with, that is, the view that the proper criteria for ethical and political choice are those given in each culture's traditions." Nussbaum is forced to acknowledge, "This universalist non-relative aspect of the view needs further development, however, if it is to prove possible to answer the legitimate worries of those who have seen all too much paternalistic imposition of some people's ways upon others." (Nussbaum 2005:5) Critics do not agree that the universalist aspect of the capabilities approach can be rescued. For instance, Charusheela (2009:1150) claims that, "I have shown that Nussbaum's approach is no more automatically safe from ethnocentric universalism, or automatically capable of addressing the hold of structural power on our discourses and social analyses, than previous renditions of modernist political liberalism."

support a particular position or set of actions than be used in order to provide a conceptual principle that may be universally applied.

The judgement approach was the final way in which proportionality might be rescued. It is incapable of providing such a rescue.

5.8. Chapter Conclusion

Proportionality is a technique that compares the significant harms and benefits that are avoided by an action with those that are caused by an action. Harms and benefits are represented by the degradation or amelioration of interests. Proportionality compares the relative status of interests in two (or more) sets of circumstances.

For proportionality to be an effective and appropriate technique, it has previously been argued that a plurality of interests must be included in this type of comparison. This is because a plurality of interests contributes to a flourishing human life. In this project, informational interests including private data control and public data access must be included in the comparison.

Espionage, and especially cyber espionage, provides salient examples of challenges to informational interests. This is because espionage is a technique that exerts direct influence on informational assets.

INFLUENCE TABLE	Direct Influence	Systems Influence
Physical Effects		Sabotage (Discrimination)
Informational Effects	Espionage (Proportionality)	Subversion (Just Cause)

Although this is not a conventional description of espionage, it was shown that it is not contradictory with previous understandings of espionage. Cyber espionage highlights aspects of the conceptual space provided by this definition in ways that conventional espionage does not. Cyber espionage is capable of manipulation of data through create and update operations to a greater extent than conventional espionage, which is usually limited to read-type operations. This does not mean that

those types of create, and update operations are not espionage. Rather, they are an aspect of espionage that becomes increasingly practical with cyber methods. Cyber methods increase the potential influence of espionage on informational assets.

The value of informational assets causes problems for proportionality. Informational interests are inherently incommensurable with physical interests to a greater extent than other physical interests. It is difficult to find a way of comparing private data control with, for example, physical injury. Many of the more conventional interests are difficult to compare. However, the informational interests, private data control and public data access, emphasise and increase this incommensurability. As the informational interests become more significant in human lives - as continues to happen in the digital revolution - the difficulty in comparison becomes more significant. There is no principled way to compare these significant interests.

If there is no principled way to compare these interests, then perhaps we might find an adequate approximation. Section 5.6 argued that such an approach would rely on our intuitive ability to make appropriate judgements between interests that were technically incommensurable. Even if the interests cannot be compared in a 'scientific' way, human judgement was able to make an adequate comparison.

It is almost impossible to disprove the fact that human judgement is sufficient to make adequate proportionality comparisons. For instance, history seems to support the idea that humans do not always make flawless proportionality judgements, but this does not prove that such judgements are impossible. Instead of a formal proof, enough reasons to doubt were amassed that the burden of proof sits firmly with those who claim human judgement is appropriate.

The reasons to doubt were that, even at small scales, there is scant evidence that our judgement is not flawed. For us to make judgements at the largest scales we would either need to appreciate the ethical concerns at that scale directly or we would need to scale-up or ethical assessments at a smaller scale to fit the larger scale. There are deep concerns about the process of scaling-up ethical assessments. The two that were briefly discussed were problems with aggregation and problems with the comprehension of the numbers involved in larger scales. The second of these concerns also impacts on the idea that we might appreciate the ethical concerns at this scale directly. If we cannot properly comprehend the numbers, then it seems unlikely that we can properly assess the harms related to those numbers.

Further, any reliance on intuition or judgement in proportionality assessments leaves the principle heavily vulnerable to challenges based on the lack of universality of values. Again, this is a subject firmly outside the scope of the project. For doubt to be cast, it is enough to believe that different people and different populations might value incommensurable interests differently.

In the presence of truly incommensurable interests proportionality cannot be rescued by the idea that our human judgement can compensate for the obvious problems.

This leaves us in the position that proportionality is inapplicable in this context and that its application may lead to inaccurate assessments.

This is an uncomfortable realisation. If we consider the foundation of the principle of proportionality as being the idea that we should avoid excessive actions, many of us would want to support that idea. How does this fit with the realisation that proportionality is not applicable in the context of incommensurable interests such as cyber interests?

One approach might be to say 'well, proportionality is the best (and only) way we have to limit excessive actions'. The thrust of this project is that with increasingly incommensurable interests this becomes increasingly problematic. Continuing to use a flawed principle creates an increased likelihood of inaccurate assessments. Moreover, it allows less benign actors space to create their own interpretations of proportionality. Specifically, proportionality allows stronger nations to justify actions based on their view of future harms. Without a coherent and universal way of assessing either the harms of the action or the harms to be prevented, such justifications are largely empty.

Another approach would be to find another principle to use in order to limit any excessive actions. It was argued earlier that proportionality is essential for a plausible moderate theory. Challenging this argument would provide an escape but proportionality embodies the idea that actions should not be excessive. Therefore, finding such a theory seems unlikely.

The final approach might be to abandon moderate theories of conflict such as JWT. It is the arrival at this conclusion that makes the realisation that proportionality is inapplicable in cyber conflict uncomfortable for many. The abandonment of moderate theories of conflict is an issue that is discussed more fully in later chapters. If it is genuinely unavoidable, what are the consequences for the ethical consideration of cyber conflict?

6 Subversion and Just Cause

> **National defense is one of the cardinal duties of a statesman, and that there is an obligation to perform such a duty absolutely irrespective of party politics or factional differences.**
>
> **John Adams, 1815**

6.1 Introduction

The brief outline of just cause that was given in the Introduction was:

> *'Just cause' is perhaps one of the most deceptively difficult of the group to pin down. In pre-theoretic terms it is easy - just cause simply means one has a worthy reason to enter into war. A more technical definition is difficult. Lazar (2020 Section 2.5) uses the phrase, 'the war is an attempt to avert the right kind of injury'. This captures the idea that a just cause must meet certain criteria, but doesn't specify what those criteria are. It is the specification of those criteria that hides the difficulties in just cause.*

In the modern formulation of JWT, and in International Relations in general, the most widely accepted provider of just cause is the defence of the state - which is most usually referred to as national defence. It is the idea of the defence of the state that will form the focus of this chapter.

This chapter is exploratory. The scope of the exploration is the effects of the frameworks that have been constructed in this project on the consideration of national defence. It is exploratory because it follows a particular path through the issue. The path it takes is persuasive, but it cannot alone provide a formal proof of the conclusions.

The path that is chosen proceeds in three stages. The first is that the authority of the state derives in some way from the members of the state and that a right to defence is inextricably tied to a duty to protect those members. Secondly, that for inherent technological reasons, the state is neither correctly positioned nor capable of protecting its members in the context of cyber conflict. Finally, tying these two points together, is the claim that the state cannot have an extant duty to protect in circumstances where it is unable to protect. If it is unable to protect, it cannot have a duty to do so. If it cannot have a duty to protect then a right of defence is implausible.

Already, it is clear that the chapter "takes seriously the idea of state-held rights".[30] National defence is interpreted as a right held by the state to defend itself. Those who disavow the idea of state-held rights may deny this interpretation but do not necessarily need to deny the conclusions of the chapter. It is possible that the argument might be refactored in more individualist terms and the denier of the conclusions would need to show that such a refactoring would alter the conclusions.

All three stages of the argument described above might be challenged. What is presented in this chapter is a plausible argument rather than a comprehensive or formal one. This line of thought is both plausible and with precedent. It is based on the compelling idea that the right of national defence stems from the fact that the state is correctly positioned, is capable, and is authorised to provide its citizens the protection they require in an antagonistic world. Following this argument, it will become clear that a right to national defence is untenable in the cyber context. At the end of the chapter, the burden will fall on a supporter of national defence to explain how this provisional conclusion can be challenged.

There is an additional distinction that is drawn in this chapter and is based on analysis of the Influence Table.

INFLUENCE TABLE	Direct Influence	Systems Influence
Physical Effects		Sabotage (Discrimination)
Informational Effects	Espionage (Proportionality)	Subversion (Just Cause)

Paradigmatic subversion in international relations is a disruption of the political system in place in a state. As such, subversion is an attack on an organisational system rather than a physical system. In the language of the Influence Table subversion is an action which has informational effects by systems influence. This reinforces the paradigmatic understanding of subversion.

[30] I use a phrase provided by David Rodin. (Rodin 2004:65)

One might question whether the state is a physical entity or an informational one. Is the state best defined by the organisational structures that exist within it? Or is the state best defined by the geographical area, resources, or citizens of the state? Or a combination of the two? Resolution of these questions is moot. What is needed for this argument is the understanding that the organisational structures[31] of the state are an intrinsic and important part of the state. Whether or not they are the only, or most important, part of the definition of the state is not relevant.

The claim is that the state has both physical and informational aspects that are intrinsic to its definition. This claim is not without precedent. It is the duality on which the bloodless invasion examples that have been discussed previously trade. In those examples, one of the questions asked is whether an invasion that caused no bloodshed or deaths might be resisted by actions that included physical harm up to and including death. In the phrasing of this project, the bloodless invasion examples pose the question of whether something of informational value (the existing organisational structures) can ever be balanced against physical lives.

Subversion is an attack on the informational aspects of the state. It does not necessarily change the geography or border of the state. It does not necessarily directly harm the members of the state. It does not destroy factories, homes, roads, or other infrastructure. It only influences the organisational structures that form an intrinsic part of the state. National defence includes the defence of the informational aspects of the state. A right to national defence encompasses a right to defend the informational aspects of the state.

Describing national defence as the defence of the informational aspects of the state is not the only way in which national defence can be, or has been, construed or formulated. Traditionally, national defence has been envisioned as a right to mobilise one's own armed forces in order to defend the territorial and political integrity of the state. For instance, the U.N. Charter emphasises prohibitions against the use of force in such cases,

> 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. (U.N. Charter 1945b)

And goes on to say,

---

[31] Organisational structures are interpreted here as non-physical. They are the schema by which things are organised. If I organise my pens by colour, the schema is that pens of similar colours are grouped together. There may also be a physical structure – say a 'desk tidy' - that I use to implement this schema. When the phrase 'organisational structure' is used in this chapter it refers to the schema, which is informational. It does not refer to the desk tidy.

> 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. (U.N. Charter 1945a)

The charter is widely interpreted as accepting a right of national defence which includes defence against threats to political independence and territorial integrity. However, those two things are significantly different. Territorial integrity is perhaps the most obvious example of the physical aspects of a state, whereas political independence may be regarded as informational. The charter accepts the duality of informational and physical aspects of the state.

As might be expected in this project, the focus of this chapter is on the informational understanding of the right to national defence, rather than the physical understanding. That is not to dismiss the physical understanding. It is simply that in the context of subversion, and in particular cyber subversion, it is the informational understanding that is central.

The consideration of national defence is made significant by the fact that extensive permissions are granted in the exercise of national defence. A state is permitted to defend itself with any amount of force, including lethal force. David Rodin expresses the scope of the permissions lucidly,

> The scope of the permission is breathtaking. States are permitted to intentionally kill enemy combatants who are deemed to be responsible (though not necessarily culpably so) for an aggressive threat posed to their sovereignty. They are moreover permitted to unintentionally but foreseeably kill non-combatants who are not responsible in any way for the threat. It is important to understand that the latter is not simply a permission to inflict harm on the innocent, but is a permission to infringe their rights, since those killed collaterally in military action have not forfeited their right not to be killed. (Rodin 2014:71)

Despite the extreme permissions that are allowed under the umbrella of national defence, the acceptance of a right to national self-defence in international law mirrors both public perception and the traditionalist standpoint in the philosophy of war. It both appears commonsensical and is widely accepted in the popular perception of international relations. Given the plausibility of a right to national defence, it would seem that its defence should be straightforward. It emerges that things are more complex. What exactly is the justification for the right of national defence? What principles ground this right? How can we explain why states have this right? Do all states have this right?

These questions will be discussed only briefly in the following sections. While they form a context for this chapter, they are somewhat moot. The reason for this is that the chapter focuses on a particular interpretation of the right to national defence. This interpretation is that the right of national defence stems from the fact that the state is correctly positioned, is capable, and is authorised to provide its citizens the protection

they require in an antagonistic world. It is not claimed that this is the only way that a right of defence, or the state itself, might be construed. Rather, it forms a prevalent thread in the popular, legal, and philosophical conceptions and is a powerful and intuitive thought that must at the very least be confronted and analysed. For ease in what follows it will be termed the intuitive conception of national defence.

This chapter argues that the societal changes involved in the digital revolution challenge the intuitive conception of national defence in two ways. The first is that, in the cyber context, the state is neither correctly positioned nor capable of doing so. The second is that in the cyber context, it is not clear that the state is authorised. As such, the state cannot meet the requirements that would be necessary of an entity that had the right to defence involving such extensive permissions.

Earlier it was stated that it will become clear that a right to national defence is untenable in the cyber context. In this project, the cyber context is defined by the logical space of the Influence Table. That implies that both cyber methods and cyber interests are within the cyber context. It is possible that the use of cyber methods, either in offence or in defence, may have effects on a putative right to national defence. That is not the focus of this chapter. The focus of this chapter is on the changes in our interests that are implicit in the digital revolution. Cyber interests are growing increasingly important in our lives.

While it is correct to say that the focus of this chapter is on cyber interests, it is also acknowledged that cyber methods have made those interests vulnerable to an extent that they have not been in the past. Interests and methods are significantly interrelated. The conceptual separation of cyber interests and cyber methods is valuable, but it should be remembered that the two influence each other in ways that make the pragmatic separation more difficult. Nonetheless, the aim of this chapter is to show that the removal of morally significant interests from national control results in a diminished right to national defence.

The conclusion will be that in the cyber context a right to national defence cannot be defended by the intuitive conception. Readers may question the scope of this conclusion. Does the conclusion challenge the right to national defence overall, or just in specific circumstances? It may be that the right to national defence does not exist in situations in which cyber interests are threatened but does exist in situations in which other interests are threatened. The aim of this chapter is to demonstrate that the state is not correctly positioned nor capable of defending its citizens' cyber interests. These issues of scope are tackled more thoroughly in the following chapter.

It was promised in the Interlude that this chapter might provide some guidance in choosing between realism and pacifism. Nonetheless, the previous chapters have argued that moderate theories fail, which leaves only pacifism or realism as available options. The pacifist is hardly committed to the idea of national defence. On the other hand, I will argue, that the realist must be committed to it. So, rather than laying out a formal argument, this chapter discusses the issues and highlights challenges to the intuitive formulation of national defence that the realist must counter if their position

is to be coherent. It aims for persuasion rather than proof. It may be that the ardent realist will be able to counter the arguments of this chapter, but whether they are willing to make the concessions that are necessary remains to be seen.

6.2 Realism and National Defence

One of the aims of this chapter is to outline a number of challenges that arise for the realist from the previous analysis of cyber conflict. This sub-section starts with a generic definition of realism based on the predomination of concern with national interests. It continues by claiming that, based on these definitions, realism implies a right to national defence. With those pieces in place, it explains how the challenges of the previous chapters provide challenges for the realist.

Common definitions of political realism emphasise national interests and state security. For instance, Duncan Bell uses the definition,

> realism, set of related theories of international relations that emphasizes the role of the state, national interest, and power in world politics. (Bell 2022)

And W. Julian Korab-Karpowicz says,

> Realism, also known as political realism, is a view of international politics that stresses its competitive and conflictual side. It is usually contrasted with idealism or liberalism, which tends to emphasize cooperation. Realists consider the principal actors in the international arena to be states, which are concerned with their own security, act in pursuit of their own national interests, and struggle for power. (Korab-Karpowicz 2018:1)

This definition is relatively weak, in that it says that ‘states are concerned’ with their own security. As Jonathan Haslam is more forceful when he says,

> Common to all realists, consciously or not, is the notion of Reasons of State: the belief that, where international relations are concerned, the interest of the state predominate over all other interests and values. (Haslam 2002:17)

Haslam goes on to defend this statement with an analysis of realist thought from Machiavelli to the modern day. For this project, the stronger interpretation is accepted: that realism implies that national interests predominate over other interests.

The aspect of the realist view that is important for this chapter is a right to national defence. This is interpreted as right that is held by the state that permits it to defend itself and its interests. The claim on which the argument hinges is that the realist position in international relations is untenable without a right to national defence. In other words, support for the idea of a right to national defence is an intrinsic part of a realist position and without it the position is not tenable. If that is the case, and it can

be shown that a right to national defence cannot be defended, then the realist position is untenable.

No greater involvement in realist theory is required. The only assumption regarding realist positions is that they require a commitment to the idea of a right to national defence. That assumption is based on the claim that all realist positions claim that national interest predominate over other interests.

It may be that this absolute requirement that national interests seems so extreme that it does not represent genuine realist views. However, a view that is not absolute does not escape the criticism that this project aims at moderate theories. A standpoint that is close to extreme realism, but still balances the costs with benefits of actions, will remain subject to the criticisms of this project. Only an extreme version of realism escapes those criticisms.

Likewise, a realist might claim that morality is simply not applicable to the arena of international relations and this talk of permissibility and rights is misplaced. As has been discussed previously, the arguments of this project will not be persuasive to the denier of the relevance of morality to large-scale conflict and international relations. Having said that, much of this chapter is not based on moral theory and the arguments that national defence is not justifiable have traction even to those who deny the relevance of morality.

The primary aim of this chapter is to show that the idea of a right to national defence cannot be justified in the scope of cyber conflict. Only a secondary aim is to outline problems that this might cause for realist standpoints. In response, realists may show how their brand of realism escapes these problems. Without clarity about how this task might be achieved, the rest of the chapter will proceed with the assumption that a right to national defence is a necessary element of realist views and this implies that any challenge to the right of national defence also is a challenge to those views.

6.3 Cyber Subversion

> **When you attack a country, it's an act of war, and so we have to make sure that there is a price to pay so that we can perhaps persuade Russians to stop this kind of attack on our very fundamentals of democracy.**
>
> **John McCain (Schleifer & Walsh 2016)**

Cyber subversion provides a paradigmatic example of cyber actions against the state itself because it is an action that has influence on informational systems of control and production.

Subversion of a state is not a physical attack. Subversion of a state is an attack on the informational aspects of the state. Cyber methods increase the reach and effectiveness of subversion and allow informational systems to be influenced

independently of physical systems. In the epigraph, John McCain makes the seemingly obvious point that “When you attack a country, it's an act of war.” What he is referring to is the action of electoral interference which he calls a fundamental of democracy. A core element of a democratic state is its electoral process and subverting that process constitutes an attack on the state.

Cyber methods provide informational vectors for attacking the electoral system of the state. There was widespread reporting that Russia interfered with the U.S. elections in 2016. The Russian techniques were widespread disinformation and ‘hack and release’ campaigns (U.S. Government 2017a). Although most of the focus is on the U.S., elections similar campaigns have targeted elections in the UK, Holland, France, and Germany to name just a few. It is likely that the Chinese exert similar influence globally, but particularly in Asia. The 2018 elections in Cambodia have come under particular scrutiny (Henderson et al. 2018). Counter-claims, that the U.S. is involved in the most widespread electoral interference, possibly have merit. It seems unlikely that the most powerful cyber power, with an extensive history of foreign political interference, is refraining from using cyber methods in pursuit of its global goals (e.g., Lange 2016). As Nick Tsagourias (2019:3-4) says,

> Although electoral interference is not a new phenomenon, cyberspace increases the scalability, reach, and effects of such interference and poses a serious threat to a State’s sovereign authority.

Nor is electoral interference the only form of cyber subversion. Informational campaigns can be, and almost constantly are, aimed at all elements of the state system. Interstate rivalries are almost invariably accompanied by informational campaigns. One of the characteristics of cyber conflict that was described earlier (§1.3.2) was its persistent nature. While physical attacks on the state are usually not persistent, informational attacks often are.

> We find that the utility of cyber strategy is as a form of political warfare optimized for the 21st century that relies on tacit bargaining and ambiguous signalling to help rival states achieve a position of relative advantage in long-term competition. (Valeriano et al. 2018:13)

Informational campaigns can target the relationship between the state and its citizens by, for instance, reducing trust in state institutions. They can target the relationship between the state and other states, for example, by influencing the status or authority of the state. They can target the system of the state itself by influencing the correct functioning of the core systems or subsystems of the structure of the state. These subsystems might include the law, social security and health systems. Anything that prevents the organisational system from working in the way it is intended can be regarded as subversion.

It has been argued in this project that traditional analyses of large-scale conflict have failed to adequately account for the value of informational assets in general. This theme is demonstrated with respect to the informational aspects of the state.

Traditional analyses have struggled to define these organisational structures of the state and their value. A particular family of examples, often described as 'bloodless invasion examples', has been leveraged to demonstrate these issues. In the context of this project, one of the issues that is highlighted by these examples is whether the informational aspects of the state can ever justify the infliction of physical harm up to and including death.

One of the ways in which the value of the state has been justified in this context is by what Michael Walzer calls the ‘common life’ of the state.

> The moral standing of any particular state depends upon the reality of the common life it protects and the extent to which the sacrifices required by that protection are willingly accepted and thought worthwhile. If no common life exists, or if the state doesn't defend the common life that does exist, its own defense may have no moral justification. (Walzer 2006a:54)

I take it that this commonality is intended as an aggregation of language, stories, history, religions, beliefs, shared moral values and the like. Whether one accepts Walzer’s particular version of a common life or not, it reflects a broader category of thought which states that individuals value the manner in which their lives are organised and how that organisation relates to their own perceived identity. It is not just their physical lives that count but something that concerns the specific qualities of citizens' lives. This echoes the emphasis that is placed on flourishing human lives that has been a thread throughout this project.

This project suggests that the organisational factors that influence individuals' lives are informational assets. In other words, what Walzer describes as a common life is a bundle of informational assets. However, one does not have to subscribe to Walzer's views or place intrinsic value on the state to believe that individuals value the way that their lives are organised. If one regards the organisation as an informational asset, then it is clear that it is highly vulnerable to cyber methods.

Remember that quote that came from the first paper to discuss cyberwar?

> We offer a distinction between what we call “netwar”—societal-level ideational conflicts waged in part through internetted modes of communication—and “cyberwar” at the military level. (Arquilla and Ronfeldt 1997:27)

A societal-level ideational conflict is exactly the type of conflict that targets the overarching idea of the way citizens' lives are organised. If national defence is justified in reference to some instantiation of this idea of organisation, it must be realised that cyber methods can target exactly that thing. The idea of netwar, and cyber subversion in particular, is crucial in the consideration of national defence in the modern world.

6.4 A Pragmatic Challenge

What has been called the intuitive concept of national defence is that the state is correctly positioned, is capable, and is authorised to provide its citizens the protection they require in an antagonistic world. One might challenge this by claiming that the world is not as antagonistic as claimed. Alternatively, one might challenge this by denying that the state has moral justification to act in this way. That is the route that is most discussed in philosophical terms.[32] An additional way to challenge the intuitive concept is to claim that the state is not correctly positioned, capable or authorised. In the domain of cyber conflict, it is this route that this chapter explores. It might be described as a pragmatic challenge to the intuitive conception and runs parallel to much of the discussion regarding just cause.

Simply because it is pragmatic does not mean that it does not hold moral significance. Pragmatic assessments of capability can have moral significance. For instance, there is extensive debate as to whether one can have a moral duty to perform an action of which one is incapable - the 'ought implies can' discussion. The 'ought implies can' discussion is active and treacherous.[33] As the Encyclopaedia Britannica says (2018), "its plausibility may depend in part on the relevant sense of 'can.'" Resolving the ought-implies-can question, or even taking a position on it is unnecessary here. What is necessary is an acceptance that capability can have moral consequences.

---

[32] This discussion is extensive. A salient example of a philosophical challenge to the right of national defence can be found in David Rodin's *War and Self-Defence* (Rodin 2002) which argues against the assumption of that right as defended in, for example, Walzer's Just and Unjust Wars (Walzer:2006a). Rodin claims that a right to national defence can be justified by two strategies and that, "The first is a reductive strategy that seeks to explain national defence as a special application of the right of personal self-defence. The second is an analogical strategy that takes seriously the idea of state-held rights and tries to account for them through an analogy with the personal right of self-defense." The reductive strategy is challenged by the fact that national defence encompasses far greater rights than are present in personal self-defence. The analogical strategy is challenged by analysing the conception of the state. If that conception is sufficiently minimal to allow all states to hold the right irrespective of their nature, then it Is hard to explain why states may defend against invaders. On the other hand, a less minimal conception of the state that incorporates a description of its nature leads to an inability to explain why one state is preferable to another. For instance, is defence permissible against an invader who will only increase the standard of life of the citizens? Rodin concludes that a right to national defence cannot be justified by these strategies.

[33] As Robert Stern says "The principle 'ought implies can' has been employed in several different debates in ethics and related areas. For example, it has been used to address the issue of free will vs. determinism; of moral dilemmas; of internalism vs. externalism as accounts of moral motivation; of obligation and blame; and of excuses and wrongdoing. None of these ways of using the principle have been entirely free of controversy, in the sense that different sides have disputed the way in which the principle has been employed to argue for one position over another. " (Stern 2004)

The claim on which this chapter is developed, that the state cannot have the right to defend itself if it is incapable of protecting its citizens, is of this type in that it relates capability and morality - although in a different manner to the 'ought implies can' discussion. The claim is plausible, and this is an exploratory chapter. It is accepted that this provides a potential avenue of challenge to the conclusions of this chapter, but the onus is on the challenger to explain how the state retains a right to do that of which it is incapable.

The reason that the onus is on those defending the right to national defence is that there are extensive permissions that accompany this right. Such permissions must not be granted without a coherent explanation. Until that explanation is provided the permissions should not be assumed and the right should be questioned.

The aim of this chapter is to show that in the context of cyber conflict the state is not correctly positioned or capable to provide protection and it will hint at the fact that it is also not necessarily authorised to do so. It is also argued that this incapability is not simply a matter of lack of resources, but rather a matter that is inherent to the state of affairs in the post-digital-revolution world.

The reason for this incapability is demonstrated most clearly by the decentralisation of citizens' interests. There are perhaps other ways in which the incapability might be demonstrated, and decentralisation is not assumed to be the only cause of incapability - but decentralisation alone provides sufficient evidence. The digital revolution has resulted in a world in which individuals' interests are not necessarily located within the state's boundaries or under the control, or purview, of the state. This process of decentralisation is continuing and results in the fact that the state is incapable of influencing morally significant aspects of individuals' lives.

6.5 Decentralisation

This section aims to demonstrate that the societal changes involved in the digital revolution challenge the idea that the state is correctly positioned, capable and authorised to protect its citizens. Only one aspect of the digital revolution, the decentralisation of cyber interests, will be leveraged although it is likely that further support might be gained from other aspects. Decentralisation provides sufficient evidence that there are significant problems for advocates of national defence in this context.

The example of decentralisation provides salient examples that challenge both the ideas of correct positioning and capability to provide protection to citizens. This is the main theme of the section. In passing, it is noted that decentralisation may also challenge the idea of the authority of the state to protect. Five aspects of the changes that are occurring due to decentralisation are highlighted in the following subsections. These are infrastructure, our personal information, the currency that we use, our work lives, and our national identity. In each of these areas there are changes that challenge the status quo and suggest that the conventional nation state no longer has

control over these aspects of our lives and will have decreasing control as the digital revolution progresses.

In each subsection an example is chosen that aims to demonstrate that the aspect discussed has immediate as well as speculative relevance. That is to say that the changes are already having effect and are not purely the domain of science fiction.

Some of the discussion may seem speculative. It is possible to dismiss certain aspects of the digital revolution as unrelated to our everyday lives. For example, do we need to be concerned with cryptocurrencies or digital autonomous organisations? In response, it is worth noting that in the 1970s it would have been hard to predict the enormous societal changes that were about to happen with the development of email, cheap global communication, and the Internet. It is plausible that the next step of that story is the development of cheap, instantaneous, and secure transfer of money on a global scale. That is the promise of cryptocurrencies. It is plausible that major financial institutions such as banks and insurance companies that have traditionally dominated the financial landscape will be transformed by distributed information and consensual software-managed contracts. This is the promise of digital autonomous organisations. While the influences on our lives cannot be predicted, it would be naïve to believe the influence will be negligible. The examples show ways in which the following discussion is already in the real world.

The purpose of demonstrating this process of decentralisation is to emphasise that the state has decreasing control over significant cyber interests of the lives of its citizens. Previous chapters have shown how significant these elements are. This chapter argues that if the state does not have control over these elements, then it cannot adequately protect its citizens.

6.5.1. Aspect One: Infrastructure

Much of the infrastructure on which the digital revolution is based is privately owned or outside the direct control of nation states.

A salient example of this is the DNS system which controls the routing of traffic on the Internet. For instance, if I want to get information from the U.S. government and know that their website is https://www.usa.gov, it is the DNS system that will ensure that my request is directed to the correct server and that the correct information is returned to me. Control of the DNS system is now organised through an organisation called ICANN and is officially independent of any particular state. The information regarding DNS records is stored on globally distributed and redundant servers.

Control of the DNS system allows traffic to be monitored and controlled. For instance, shutting down access to global DNS servers from particular countries can seriously degrade, to the point of non-functionality, the use of the Internet in those countries. ICANN was originally founded in the U.S. but is now overtly independent of any state structures or influence. Currently, the tagline on the ICANN website reads, 'One World, One Internet'.

It is problematic for states that an essential element of information routing is outside their control. The answer to this lack of control results in plans by more authoritarian states to regain control by constructing domestic alternatives to the current DNS systems (e.g., Epifanova 2020). Such moves would fracture the idea of a universal Internet but are defended by the argument that most states do not have control of information routing themselves. The idea that the DNS system can be manipulated is alarming for anyone wishing to control public data access, either for benign reasons or oppressive reasons. While a state does not control the DNS system, it cannot claim to control the information that its citizens receive. Conversely, if they aim to control this information routing, they are open to charges of censorship and information oppression.

At the time that this section is being edited, the Ukrainian Government is asking ICANN to remove Russian domain names from the global DNS servers. This is the first time that control of the DNS system has been overtly considered as an implement of state action or international influence. Eamon Javers (Javers 2020) says, "The move, which would be unprecedented, could spur Russia, China, and other governments to move away from ICANN and spur the balkanization of the internet."

It is clear that control of the DNS system is impossible for an individual country – even the U.S. - and that this creates challenges for a state wishing to control information flow to and from its citizens, either for benign or sinister reasons. In this respect, states' abilities to protect their citizens' cyber-interests against attacks are minimised. The DNS system is only a single example of this process of global infrastructure being removed from state control. The Internet backbone – the actual major physical cable and associated equipment and functions – is privately owned. It might be argued that the major Internet software that is used – WeChat, Facebook, Twitter, etc. - also fits this definition of being outside the control of the state. In fact, more challenging is to imagine an element of the infrastructure of the digital revolution over which the state does have significant control.

The increasing lack of state control over the infrastructure on which the digital revolution is based demonstrates that these problems are a factual rather than speculative matter.

6.5.2. Aspect Two: Personal Information

Two families of digital interests have been developed in this project, private data control and public data access. The influence of the state on both these families is being challenged.

Regarding private data control, the landscape is changing in that as part of the digital revolution, all the information concerning an individual is being digitised (e.g., N.H.S. 2022). It seems likely that in the near future the vast majority of information regarding individuals will be stored in digital form. This applies to birth and death records, health

records, financial records, work and employment records, tax records and many more.

It is completely possible for the state to be the guardian of all this information in a secure and ethical manner. However, there are reasons to believe that is not likely to be the case. Firstly, little is absolutely secure. Secondly, the investment involved in doing so is immense and the zeitgeist in many countries is to outsource these functions. Thirdly, in many countries, there is neither the experience nor the resources to implement such a system. Fourthly, there is strong persuasion from outside actors to allow access. All these reasons, and more, suggest that increasingly the information that is important in the provision of flourishing lives will be outside the direct control of the state. It is more likely to be held by multinational corporations.

Moreover, with the advent of widespread redundant backups and decentralised storage, the very concept of geographic location is becoming obsolete in regard to data storage. Decentralisation means that the data is stored consensually across the network rather than in one single place. For a system such as a geographically located state, this decentralised data storage is becoming challenging since its members' significant data is not stored within the geographical boundaries of the state nor is it necessarily under the control of state institutions. Control of private data is an interest that is increasingly difficult for a geographically located state to assure its citizens.

In terms of public data access, similar problems exist. In the past, it might have been possible to claim that the state had an adequate degree of control over the public access to reliable and truthful information. That is becoming an increasingly implausible claim.

The most topical example of this issue is electoral interference. It is widely accepted that there was significant Russian influence exerted in the U.S. Presidential elections in 2016. The method of influence was not physical but was a widespread and effective information campaign. The U.S. is undeniably the most powerful entity in the cyber domain, but it was unable to assure that the information that its citizens received was truthful, unbiased, and reliable. The problem stems from the fact that the information pathways by which citizens receive their information are no longer controlled by the state.

This example highlights a theme that will reoccur in this chapter. There are ways in which this problem can be combated but they involve draconian measures. For instance, the Great Firewall of China is an attempt by the Chinese state to exert control. In the democratic world, these measures are widely perceived as totalitarian and infringing civil liberties and freedom of information. The idea of net neutrality stands at the centre of a liberal conception of the Internet. It demands that those entities that are involved in the infrastructure and applications that convey information are neutral to the information that is conveyed. The fact that even democratic states

often oppose net neutrality reflects the difficulty that freedom of information transfer causes the state.[34]

State control over both private data control for its citizens and public data access for its citizens is already degraded by the changes of the digital revolution. The state's correct positioning and capability to protect the cyber interests of its citizens is reduced.

6.5.3. Aspect Three: Currency

Traditionally, one of the cornerstones of the architecture of a state is the provision of a national currency. This cornerstone is being attacked by the ongoing development of cryptocurrencies that are based on the related technologies of distributed ledger technology (DLT) and blockchain.

The technologies are able to provide secure and synchronised digital data that is stored across geographically distributed sites. There is no centralised administrator and the actions on the network are consensual and distributed. This provides a secure method of data storage that is entirely decentralised. An implementation of blockchain that is discussed in this subsection is cryptocurrency. Another implementation, smart contracts, underpins the next subsection.

The success of bitcoin is the best-known example of the actual use of a blockchain. Bitcoin is not alone, and the overall investment and adoption of cryptocurrency are growing. The reasons for this are various.

> The heightened distrust in the traditional financial system and the exclusive power that financial institutions have over consumer funds are just some of the reasons. There's also the high inflation rates, unstable economies, and lack of privacy. Not to mention the impractical cross-border payment system. (Page 2021)

Cryptocurrencies can solve problems associated with conventional currencies. Money can be moved cheaply around the world, smart contracts (discussed below) can facilitate entrepreneurship and business deals that would otherwise be impossible, and micropayments can be made effectively. Cryptocurrencies can be seen as a protection against high inflation in countries with struggling economies. In 2017 Turkey had the highest adoption of cryptocurrencies for this reason (Das 2018). Cryptocurrencies can provide a level of privacy concerning financial dealings that many users desire.

---

[34] See, for example, Morton 2019 for the struggle for net neutrality against the U.S. federal legislation that undermined that neutrality.

Cryptocurrencies move the control of finances away from the traditional control of commercial and national banks. For this reason, it is a move towards decentralisation that is being fought vigorously by existing powers. The fight is taking place both in the media and in the courts. In the popular media, cryptocurrency is often painted as the domain of criminals and tax avoiders. Cryptocurrency losses receive a great deal more coverage than crypto gains. If one buys cryptocurrency with a UK bank account one will be warned that one might lose all one's money - a warning that would not be forthcoming on almost any other investment. In fact, although the cryptocurrency market is extremely volatile, long-term investments have followed an upward curve; a fact that explains the increasing involvement of mainstream corporate investment in the crypto market. In the courts, various legislation has been introduced around the world to try and limit cryptocurrency. Measures in China have been extreme (Sigalos 2021), and the recent Infrastructure Bill in the U.S. (Reitman 2021) is another example of the fight that is underway globally to limit cryptocurrency adoption. The struggle against the decentralisation of finances is one of the few areas in which China and the U.S. are aligned.

The Internet has transformed the way that information is moved around the world. Effectively, money is simply data, so it seems unlikely that the corresponding changes will not occur in financial matters. The value in moving money around the world in a frictionless manner is simply too great for cryptocurrencies not to be adopted. It may be that the only hope for the national banks is to release their own cryptocurrencies. China is potentially leading the way with the digital Yuan (Kharpal 2021). Whether this type of national digital currency can offer the advantages of cryptocurrencies remains to be seen.

It seems unlikely that states will be able to retain absolute control of the currency by which citizens judge their wealth and correspondingly assess their welfare. This fact challenges the state's ability to protect its citizens' interests.

6.5.4 Aspect Four: Work Life

The changes of the digital resolution are altering the nature of the businesses that form an important constituent part of our lives and welfare.

The idea of decentralised finance (DeFi) in which conventional financial services are offered through blockchain solutions is growing (Anissimov 2021). Blockchain technology allows the implementation of 'smart contracts' which are contracts that exist entirely on a blockchain. An example of a smart contract might be a system in which farmers contributed payments into a 'kitty' and were paid from the kitty in cases of drought. The criteria of payment into the kitty and reward from the kitty exist only in a software contract. As Alexander Savelyev (2016) suggests in his paper, *Contract Law 2.0: «Smart» Contracts As the Beginning of the End of Classic Contract Law*, such a contract would replace the contract the farmers conventionally would have with insurance companies. The legal standing of smart contracts relative to traditional contracts is actively debated (e.g., UK Government 2021). There are great

advantages to financial services outside the scope of government regulation. There are correspondingly large risks associated with the lack of regulation. Again, how that tension is resolved remains to be seen. Smart contracts remove the control of financial contracts such as insurance contracts from the control of the traditional institutions.

Decentralised finance also drives the increasing development of DAOs, Decentralised Autonomous Organisations. DAOs are networks which allow people in different locations to collaborate to fulfil a particular project or aim by using cryptocurrencies to facilitate these projects.

> The thing is, not every problem is a big one, and not every task requires a large group of people. An underrated aspect of crypto networks is their ability to enable cooperation on a small scale, but amongst people spread across the globe, living in varied legal jurisdictions. The experimentation with DAOs that we're seeing on Ethereum is an exciting and instructive example of this.
>
> Even in this germinal stage, what's happening with Moloch and MetaCartel is quite remarkable when you stop to think about it. Dozens of individuals, from many different countries, are pooling millions of dollars to fund public goods in a digital community they care about. They're doing this without an official legal structure or legal documents, and without any centralized executor or escrow service. A decentralized network, and the smart contract code that runs on it, is all that is needed. (DiFrancesco 2019)

Needless to say, the conventional finance sectors and associated political organisations do not view this favourably and are fighting back both in legislation and in rhetoric. The legislation included a 2017 ruling by the United States Securities and Exchange Commission (SEC) that was aimed at making involvement in DAOs considerably more difficult, if not illegal (U.S. Government 2017b). The rhetoric is based around the insecurity and risk of cryptocurrencies in general and the lack of regulation that is available for DAOs.

The move towards decentralisation of financial and business entities is likely to continue. It is driven by the ability of blockchain technology to provide secure and easily accessible data in the form of DLTs that is largely out of the control of external control or regulation. Increasingly, the state is unlikely to be in control of how we organise our working lives, the business entities that we construct, the entities that provide employment, and the way that we understand our employment.

Again, these factors will minimise the state's ability to protect these important areas of citizens' lives.

6.5.5. Aspect Five: National Identity

A symptom of the decreasing link between citizens' work and their geographic position is the growing popularity of e-residency for business purposes. The best-known

example of this is e-residency in Estonia. Other countries that have e-residency programs are Lithuania and Ukraine, with others reported as considering the option. The idea on which e-residency is based is that it is possible to set up a business under the jurisdiction of that country while remaining a national of another country. There is no requirement that one is present in, for instance, Estonia to do this. Estonia allows individuals and corporations to operate as Estonian companies while having no physical presence in the country.

> Estonia is the first country to offer e-Residency, a government-issued digital identity and status that provides access to Estonia's transparent business environment: a new digital nation for the world. (Government of Estonia 2022)

The advantages of e-residency in Estonia are access to the European market, a stable political situation, and a favourable tax environment. E-residency gives no right to physical residency, nor does it allow access to state institutions such as health care. It is possible for an individual to conduct their business life entirely under the jurisdiction of Estonia while being a citizen of another country.

The implications of e-residency are complex. For the purposes of this very brief summary, the points that are important are that the individual's home state no longer has control over the financial life of that individual and is unable to protect that financial life. As a side note, of course, the idea of e-residency also moves towards a devaluation of the affiliation of the individual to the state and therefore potentially reduces the affiliation of the individual to the state.

Again, how the concept of e-residency develops remains to be seen. Nonetheless, it is a clear example of how a state may lose control over the lives of its citizens, how the state may lose the ability to protect citizens and how individuals may lose the strong affiliation to the state on which the state's authority is ultimately based.

6.6.6. Are These Claims Contentious?

The main claim of the section is that the state has decreasing control over significant aspects of the lives of its citizens. One way in which this might be countered is to claim that although the examples above are valid, the ultimate control still resides with the state. So, for example, whoever owns the infrastructure of the Internet, it is the state that has ultimate control within its geographic domain.

One might expect states to adopt this perspective. However, there are areas in which the language of the discussion suggests that even the strongest advocates of state control are forced to accept that the state cannot be the ultimate authority in certain areas. This is most clearly seen in the international framing of the development of cyber norms in terms of the multi-stakeholder model.

The multi-stakeholder model is one in which it is implicitly accepted that not only state authorities are capable or responsible for the development of Internet norms. In its simplest form, the multistakeholder model implies that various entities, states,

corporations, user groups, and other institutions all have relevance and authority in the field.

The U.N. discusses the generation of cyber norms in terms of the multi-stakeholder model. A discussion in 2020, titled "Operationalizing Cyber Norms", included the description (United Nations 2020), "This panel discussion, the first of a series of multi-stakeholder dialogues on cyber norms, will bring together representatives from different stakeholder communities to discuss challenges and opportunities of operationalizing this norm." Shears et al. (2018) start their paper on approaches to cybersecurity with the claim, "For some time, governments and non-governmental actors alike have been calling for greater stakeholder involvement in cybersecurity policy." Another example is that the mission of the Global Commission on the Stability of Cyberspace, launched at the 2017 Munich Security Conference, released its final report in 2019 in which it claimed,

> The Commission concluded its report with six recommendations that underline the need for strengthening the multistakeholder model. (Faesen & Roggeman 2019)

The idea of a multi-stakeholder model is widely embedded in the discussion regarding cybersecurity. It is often justified, as above, by an inclusive policy in which expertise is gathered from those who have it. However, it also implies an implicit acceptance of the fact that state governments are not correctly positioned or capable of managing cybersecurity alone.

The fact is that the multi-stakeholder model contains an implicit acceptance that the state is not in control of these aspects of the post-digital-revolution world. This chapter works with the idea that the state has a decreasing influence on aspects of its citizens' lives. The multi-stakeholder model is an acceptance of this fact.

6.5.7. Decentralisation: Conclusion

It has been demonstrated that the information that is essential for providing flourishing lives is outside governmental control and that our work and financial lives are increasingly outside the control of the state. Five key areas were used as salient examples of this fact, the infrastructure of the digital revolution, the storage and access of personal information, currencies, the structure of our work lives, and our concept of national identity. It is likely that this process, one of distancing important aspects of lives from state control, is likely to escalate as the digital revolution progresses. The prevalence of the multi-stakeholder model shows a widespread acceptance of this fact.

The result of this distancing of significant aspects of our lives from government control has two consequences. Firstly, state governments are increasingly unable to provide the protection that is required by their citizens. Secondly, because of this first consequence and because of an increase in external affiliations, the affiliation of citizens to the state is diminished. As was stated earlier, how this second

consequence affects the authority of the state is open to extensive philosophical discussion and is beyond the scope here, but it seems unlikely that it can have no effect.

6.6 Conclusion

Most discussion surrounding the right to national defence focuses on the ethical justification of the right itself. In this, this chapter has taken an unusual route. It has challenged the ability of the state to be correctly positioned or capable of providing the requisite protections. That might be considered a pragmatic challenge, but the context gives this pragmatism moral significance.

It is worth restating the overall form of the argument that has been explored. Its three steps were outlined in the introduction.

> The first is that the authority of the state derives in some way from the members of the state and that a right to defence is inextricably tied to a duty to protect those members. Secondly, that for inherent technological reasons, the state is neither correctly positioned nor capable of protecting its members in the context of cyber conflict. Finally, tying these two points together, is the claim that the state cannot have an extant duty to protect in circumstances where it is unable to protect. If it is unable to protect, it cannot have a duty to do so. If it cannot have a duty to protect then a right of defence is implausible.

The work of this chapter was to defend the second step, to demonstrate that the state was not correctly positioned or capable of providing protection to citizens. The chapter was exploratory. In this it did not fully develop the first and third steps. Rather, it presented a challenge to those who might desire to defend the right to national defence. Given that Step 2 has been plausibly demonstrated, an advocate of the right of national defence will need to challenge the other steps in the argument. In other words, it will be necessary to show that the authority of the state does not derive from the members. Or it will be necessary to show that an intrinsic incapability to defend does not make a right to defend implausible.

In order to demonstrate Step 2, the factor that was leveraged was the decentralisation that is inherent in the digital revolution and the societal changes that it has created. Decentralisation provides salient examples that are powerful enough to support the aims of this exploration. The infrastructure of the digital revolution is not under state control (§6.6.1). Personal information that is significant in citizens' lives is not under state control (§6.6.2). The currencies that dictate the lives of citizens' are decreasingly under state control (§6.6.3). Changes of the digital revolution are potentially causing changes in our relationship with finance and employment which reduce the control of the state (§6.6.4). There are changes to the idea of nationality that reduce state control in this area (§6.6.5).

The inevitable conclusion is that the state is losing control over significant aspects of citizens' lives. This conclusion might be cashed out in a number of ways. For example, those who talk in terms of government as a coercive entity will see challenges for such governments in this lack of control. The aspect of the conclusion that is leveraged here is that a state that is not correctly positioned nor capable of providing protection struggles to claim that it has a right to defend itself. The right of national defence becomes increasingly untenable as the control of important aspects of citizens' lives moves away from state control. This is because the right of defence is inextricably linked to a duty to protect these citizens. Without control, protection cannot be assured. If protection cannot be assured, then a claim to a right of national defence is not tenable.

This chapter was not intended to provide an unchallengeable argument. Rather, it was intended to highlight difficulties for advocates of national defence in the context of large-scale cyber conflict. There are a number of routes that those advocates might take in avoiding the conclusion that national defence cannot be justified in this context.

Firstly, the initial assumption might be challenged. The argument has proceeded from what was termed an intuitive justification of national defence, that national defence is grounded in the fact that the state is correctly positioned or capable of providing protection to its citizens. A conception of the relationship between the state and its members that did not demand protection of the members of the state seems implausible. More accurately, a state that did not aim to protect its members is usually considered to have significantly diminished authority. It is not clear how the assumption might be plausibly challenged. As David Rodin notes, if an extremely minimal conception of the state is adopted – one in which protection of members is absent, for instance – the difficulty that emerges is defining why such an entity might hold a right to defend itself.

Secondly, perhaps the state really is capable in some way of protecting its citizens. Recent developments suggest that this is not the case, as has been discussed, but the discussion exists in a technological domain and further developments and technologies are possible. The challenge for states is to provide adequate protection while allowing adequate freedoms. It seems unlikely that a technological solution to this age-old problem is likely to emerge in the near future.

Thirdly, perhaps rights are not defined by capability. A realist might argue that the state can maintain a right to national defence even if it is evidently incapable of fulfilling the demands of national defence. At this point, the concept of a right to national defence is so devalued, so much a shadow of what is usually meant, that it ceases to be meaningful.

Fourthly, perhaps the example of decentralisation does not prove the point that has been claimed. The point, in summary, is that the state is losing control over significant aspects of citizens' lives. A strong counterargument to that, or a strong denial, may be possible but it is not immediately evident. It would be difficult to argue that decentralisation was not occurring due to the digital revolution. After all, one of the

pillars of that revolution is the provision of near instantaneous information across the globe. This information transfer does not respect geographic boundaries and it is this fact that underpins the process of decentralisation.

So far, it seems like it is an uphill path for the advocate of national defence and that path is only likely to get steeper. There is a final challenge which is perhaps more interesting. The heart of the challenge is the idea that the argument only applies to cyber conflict. That in other domains, the conventional right to national defence still exists and is unchanged. This challenge is misguided for two reasons; it misinterprets the way that cyber has been used throughout this project and misunderstands the growing importance of cyber interests in people's lives.

It misinterprets the way 'cyber' has been used throughout this project because it posits a separation between cyber interests and human interests. Cyber has been defined as referring to the relationship between humans and information. The misinterpretation stems from an idea that the conclusion only refers to 'computer stuff.' In fact, the conclusion refers to informational factors that are significant in the provision of flourishing human lives.

It misunderstands the growing importance of cyber interests in people's lives because it intimates that protection of an individual is possible without protecting that person's cyber assets. That is incorrect already. Its failure as a standpoint will become increasingly obvious as our cyber interests continue to increase in significance in the provision of flourishing.

Nevertheless, the argument that the conclusion only applies to cyber conflict, or the use of cyber methods, is intuitive to many people. The idea is that just because the state cannot provide particular protections does not mean that it cannot be justified in providing other protections. The claim might be that the traditional right of defence that is held by the state is not affected by the digital revolution. To defend this claim, one would have to explain why, when almost every other aspect of life is affected, this right is impervious to the digital revolution. However, the scope of this conclusion, and the overall conclusion of the project, are tackled more fully in the final chapter.

This chapter was exploratory, so absolute conclusions are not appropriate. Rather, it is clear that the digital revolution creates extensive challenges for those attempting to justify a right to national defence. It may be that those challenges can be answered, but until that is the case, any claimed right of defence must be regarded as dubious at best.

7. Conclusion

7.1 Overview

The digital revolution is effecting deep and significant societal changes. It is inevitable that those changes would extend to the reasons and methods that we use to engage in conflict. Conflict, in general, has moral significance because of the harm that it is likely to create. Large-scale conflict is the type of conflict that imposes harm on groups or populations rather than just on the individual. It is the moral significance of large-scale conflict and its ability to inflict widespread harm that has motivated this project.

The arena of the project has been cyber conflict. This has been defined by use of the Influence Table that has been developed to demarcate types of cyber actions. The Influence Table defines a logical space in which cyber actions can be distinguished. It also makes sense of, or adds a conceptual grounding, to a widely supported view that cyber actions are either sabotage, espionage, or subversion. Most importantly for this project, it also defines the scope of what is referred to as the cyber context. This definition of the cyber context is pragmatic and useful, and includes an emphasis on both the use of cyber methods and the targeting of cyber interests. Cyber conflict is defined as conflict in this context. It is the focus on both cyber methods and interests that makes this definition of cyber conflict both unusual and more complete than many understanding of the terms that focus primarily on cyber methods.

The project's aim is to investigate whether existing instances of moderate war theories such as Just War Theory could be applied successfully to adapt to societal changes of the digital revolution. It was shown that the changes of the digital revolution challenge two essential elements of a theory of war, proportionality, and discrimination. In the cyber context, those principles cannot be applied in a way that creates meaningful results. Without these principles, a moderate theory is not plausible. A theory that allowed excessive actions would not be plausible. A theory that allowed indiscriminate actions would not be plausible

At this point, it might seem that the aim of the project has been achieved. Moderate theories cannot be successfully applied in the context of cyber conflict. However, a legitimate question is what this means for philosophical and ethical treatments of large-scale conflict. Should we ‘outlaw war’ (Hathaway & Shapiro 2017) as was attempted by political activists in the 20th century? Or should we defend the right of states and other groups and organisations to pursue their ambitions irrespective of the harm created? Should we become pacifists or realists?

Without acknowledging that question the project could seem arbitrary or empty. The tools and frameworks that are developed throughout the project do have something to say about these questions. It is beyond the scope to fully develop those things, but in Chapter 6, the question has been acknowledged and the upshot is that the changes of the digital revolution do not appear to create additional hurdles in the pacifist route whereas they do create additional hurdles in the realist route. The realist route appears to become increasingly difficult to justify in the cyber context as the digital revolution progresses. Any claims beyond that general observation will have to wait for later investigation.

The claim that moderate theories cannot be successfully applied in the cyber context is a strong claim and it is worth revisiting the arguments and assumptions that led us here.

7.2 Framing the Question - Part 1

This project has differentiated between cyber methods and cyber interests and has insisted that ‘cyber’ is not independent of reality or humanity. In fact, cyber interests can only be defined in relation to humanity. The term ‘cyber’ had been used to denote the critical relationship between humans and the informational assets or methods that affect them.

It is a critical relationship because of the increasing importance of these factors in providing legitimate and flourishing human lives. While simply being alive is clearly important, it alone is not sufficient to provide a meaningful and valuable human life. Informational interests have always had significance in our lives but with the digital revolution, this importance has increased and continues to increase.

While the effect of cyber methods has received significant attention, the effect of cyber interests on theories of war has been largely neglected. The fact that our interests are changing will inevitably change the reasons that we choose to engage in large-scale conflict and will require our existing ethical frameworks in this context to be re-evaluated.

In Part 1, the cyber interests framework is developed that uses a pluralistic assessment of human interests or welfare. This framework includes informational or cyber interests that are significant in an individual's life. Although it is acknowledged that those interests are very much still developing, two families of interest are proposed: private data control and public data access. Both those families are essential for a fully flourishing human life.

The aim of the project can therefore be defined as analysing moderate theories of war when applied to cyber conflict using the tools of the cyber interests framework. Moderate theories of war are those theories that aim to balance the harm of actions against the benefits of those actions. They are an intuitive approach because they

neither deny the harms nor the benefits. As such, they have formed the most widely used framework for the moral assessment of large-scale conflict.

Given the changes inherent in the digital revolution and the high moral cost of large-scale conflict, it is essential that the legitimacy and effectiveness of moderate theories of war are continually monitored. The question that emerges is whether moderate theories can adapt to a world in which cyber interests and methods are increasingly significant.

7.3 Answering the Question – Part 2

Three pre-theoretic demands of a successful moderate theory were outlined. The first is that the harm of any action should be directed appropriately. The harm should not be indiscriminate. The second is that the level of harm should be appropriate to the benefit anticipated. The harm should not be excessive. The third is that there should be a good reason for the actions. That there should be a worthy objective. In more technical terms, these pre-theoretic thoughts equate to the principles of discrimination, proportionality and just cause.

It was not claimed that these were the only appropriate pre-theoretic demands or technical principles that might be applied to moderate theories of conflict. What was claimed was that implementations of those pre-theoretic thoughts are essential for a successful moderate theory of war. This assumption is based on the understanding that a theory that allowed indiscriminate harm could not be successful, a theory that allowed excessive harm could not be successful, and a theory that allowed harm for no good reason would be unsuccessful. The challenge of Part Two is to show that in the cyber context the principles of discrimination and proportionality fail. The principle of just cause may not entirely fail in its broadest conception, but the effects of the digital revolution are such that its practical implementation becomes dubious.

Thomas Rid rather famously claimed that cyber methods were best categorised as sabotage, espionage, or subversion.[1] In this project, it is not assumed that this means that cyber methods are of lesser significance than conventional methods. To understand the significance of cyber methods, and explain the reasons behind Rid's categorisation, the distinction between direct influence and systems influence was introduced. Actions that have system influence do not necessarily affect an interest itself but rather affect the systems of control or manufacture of that interest. In the physical domain, a clear example of that is sabotage. Sabotage may affect the production of shoes but does not target the shoes themselves. An act of cyber sabotage may attack the control systems of an industrial process, for example.

Another distinction that has been used is whether the action has its ultimate effect on physical or informational assets. Sabotage has physical effects. Even examples of cyber sabotage, such as Stuxnet[2], have physical aims. In the case of Stuxnet, these aims were the destruction of centrifuges in the industrial complex at Natanz. However,

the target of an act may also be informational. For instance, in a paradigmatic example of cyber espionage, the target is data.

Combining the two distinctions that are outlined in the paragraphs above creates the Influence Table that has been used extensively in Part 2. The context of the Influence Tabe is cyber conflict and that implies that the actions that are described are harm-causing in their intent and nature.

INFLUENCE TABLE	Direct Influence	Systems Influence
Physical Effects		Sabotage (Discrimination)
Informational Effects	Espionage (Proportionality)	Subversion (Just Cause)

The Influence Table explains why the categories that were described by Thomas Rid are inevitable. It also allows discussion of these factors in a principled way because it defines a complete logical space. All cyber methods occur in this logical space and so we can be assured that no category of method is overlooked. Moreover, the Influence table also relates to cyber interests by discussing informational effects and so avoids the potential pitfall of assuming that only physical interests are of importance. Happily, the Influence Table also has three elements that are relevant and these form appropriate pairs with sabotage, espionage, and subversion. The pairings are fortuitous rather than significant or exclusive. It should not be taken that proportionality does not apply to subversion or sabotage, for instance.

The Influence Table provides a useful tool for analysis of cyber actions. The fact that it is created from foundational principles and matches the observations of Thomas Rid and others demonstrates its relevance. It gives a conceptual understanding of the categories that have been empirically observed. It allows the concepts of sabotage, espionage, and subversion to be analysed in a context that is based on foundational principles. It is a valuable conceptual tool.

In Chapter 4, the epistemic challenges of discrimination were targeted. The epistemic problem of recognition focuses on whether we can distinguish between those who meet a supposed criterion of discrimination. The epistemic problem of accuracy focuses on whether we are able to apply our intended actions to the correct group

while avoiding collateral damage. The chapter demonstrated that neither of these problems could be surmounted in the cyber context. Even if we were able to agree on which group of people were liable to harm (which seems unlikely in itself) the epistemic problems render any application of the principle untenable. This fact is due to inherent characteristics of cyber conflict rather than any transitory uncertainty or 'fog of war'. The principle of discrimination is not applicable in a principled way in this context.

In Chapter 5, the principle of proportionality was the target. The main theme of the chapter was the incommensurability problem. The value of informational assets and interests was not comparable in a meaningful way to physical assets and interests. Without comparison being possible the principle resulted in results that were at best indeterminate, and at worse simply inaccurate. Because informational assets are critical in the provision of meaningful and flourishing lives, they cannot simply be ignored in proportionality assessments, and this leaves us in the situation where proportionality is not a useful or accurate tool. Proportionality cannot be applied successfully in the cyber context.

Without effective and functional instantiations of the principle of discrimination and the principle of proportionality, a theory of large-scale conflict cannot be plausible. A successful theory must prohibit excessive actions and must prohibit indiscriminate actions. In the cyber context, Just War Theory and other moderate theories of large-scale conflict are not plausible because fundamental elements of those theories are either inapplicable or faulty. Faulty should be taken here as meaning that they return results that are either indeterminate, confused, or wrong.

7.4 Further Investigations

The aim of the project was to analyse whether moderate theories were able to cope with the changes that are implicit in cyber conflict. Chapters 4 and 5 provided the nails in the coffin for moderate theories in this context. But where does that leave us?

If one removes moderate theories the remaining options are extreme, either a fundamental pacifism or a fundamental philosophical realism. This assumes that there is a requirement for a standpoint that is defended on philosophical and ethical grounds. Without this requirement, one can think what one likes. One might even deny that ethics is applicable to large-scale conflict. With the requirement, one is forced to extreme pacifism or realism in a philosophical form. Neither is without its challenges.

Chapter Six aimed to provide some commentary on the decision between pacifism and realism using the tools that had previously been developed. The particular focus was the idea of national defence which was taken as a necessary commitment for a realist position. It was argued that there is some evidence that a right to national defence will be increasingly difficult to justify as the digital revolution progresses.

In the simplest of terms this difficulty can be understood by the fact that while states are geographically defined, our interests are becoming increasingly decentralised. The result of this process is that many of our important interests are no longer under the control of the state. The line of argument that was applied in the chapter was that if a state was not able to provide protection to its members, then the right it had to defend itself is challenged.

This chapter was wholeheartedly exploratory in that there were no hard and fast conclusions, no absolute proofs, and only empirical observations regarding the increasing difficulty for those of realist persuasion. Despite this the value of the frameworks that were previously developed was obvious and there is clearly much more that can be done in applying them in a rigorous fashion to other areas of the study of theories of large-scale conflict.

7.5 The Scope of this Result

Any dismissal of moderate theories of large-scale conflict is a strong and alarming conclusion. A legitimate question is what the scope of this conclusion is. Does the conclusion imply that moderate theories cannot be used only in scenarios in which cyber methods are used? Or those in which cyber interests are targeted? Or does it cast doubt on the use of moderate theories overall? What is the context of this conclusion?

Throughout this project, the phrase ‘the cyber context’ has been used but there are options as to what that might mean and correspondingly what the significance of the conclusion is. The narrowest interpretation of the cyber context is to claim that it is defined by scenarios in which traditionally defined cyber methods are used. That perhaps is the most widely understood definition of cyber conflict. The benefit of using this definition of the cyber context would be that it limits the uncomfortable conclusions of this project. One might be able to claim that moderate war theories were only inapplicable in a very narrow set of circumstances. Additionally, one might be able to claim that all this theorising had a trivial effect on how moderate theories should be applied to conflict overall.

That narrow interpretation of cyber conflict is opposed to the way that cyber conflict has been understood in this project. Cyber conflict has been defined both in terms of cyber methods and cyber interests. The digital revolution has changed both the way we fight and the reason that we fight. Considering only one of these provides an incomplete picture of cyber conflict.

A more complete framework has been developed that includes both cyber interests and cyber methods. It has been demonstrated that this framework more adequately describes and explains the characteristics of cyber conflict. The Influence Table not only describes cyber sabotage, espionage, and subversion, but explains why these are the categorisations that are significant. It relates also to both physical and

informational assets. It defines a coherent and complete logical space for cyber conflict that is not lacking in the manner of the narrow definition described above.

The narrow definition should be dismissed. At the very least, the scope of the 'cyber context' should be defined as occurring in the logical space defined by the Influence Table.

INFLUENCE TABLE	Direct Influence	Systems Influence
Physical Effects		Sabotage (Discrimination)
Informational Effects	Espionage (Proportionality)	Subversion (Just Cause)

The phrase 'at the very least' was used carefully in the previous sentence. That is the extent of the claim of this project -that the cyber context is defined by the logical space of the Influence Table and that in this logical space moderate theories cannot be applied in a principled manner.

However, some readers might wonder if the conclusions have even broader application than this. The Influence Table is not only applicable in 'cyber' terms. The table is, in the vocabulary used in Part 1, transversal, in that it relates to both physical and informational interests. What about sabotage that is carried out using physical means? Would it be possible to construct an argument that was parallel to Chapter 4 for these cases? Even though this is outside the scope of this project it does not seem inconceivable. Likewise, would it be possible to construct an argument that mirrored Chapter 5 in the physical domain? After all, not only cyber harms raise the issue of incommensurability. No two harms are commensurable in the strictest terms. This fact allows the possibility that the conclusions of this project might also have traction in conventional conflict. It seems possible that cyber conflict is providing salient and effective examples of problems that exist in the fabric of discrimination and proportionality independent of the context. Whether that is the case or not is left for further analysis.

The conclusion of this project is that in the cyber contexts, as defined by the Influence Table, moderate theories cannot be applied in a principled manner to large-scale

conflict. This is given moral significance because that large-scale conflict inevitably involves the infliction of widespread harm.

8. References

Adams, S. (2013). *Your confidential medical records for sale... at just £1*. Mail Online. https://www.dailymail.co.uk/news/article-2396362/Your-confidential-medical-records-sale--just-1-Hunt-insists-plan-sell-details-private-firms-vital-combat-epidemics--critics-fear-unprecedented-privacy-threat.html

Adams, T. (2012, July 7). This much I know: Daniel Kahneman. *The Guardian*. https://www.theguardian.com/science/2012/jul/08/this-much-i-know-daniel-kahneman

Albright, D., Brannan, P., & Walrond, C. (2010). *Did Stuxnet Take Out 1,000 Centrifuges at the Natanz Enrichment Plant? | Institute for Science and International Security* [Text]. https://isis-online.org/isis-reports/detail/did-stuxnet-take-out-1000-centrifuges-at-the-natanz-enrichment-plant/

Anissimov, K. (2021). *The Continuing DeFi Growth of 2021 | Finance Magnates*. Financial and Business News | Finance Magnates. https://www.financemagnates.com/cryptocurrency/the-continuing-defi-growth-of-2021/

Aquinas, T. (2006). *Summa Theologica, Part II-II (Secunda Secundae) Translated by Fathers of the English Dominican Province*. https://www.gutenberg.org/ebooks/18755

Arquilla, J., & Ronfeldt, D. (1993). Cyberwar is Coming! *Comparative Strategy*, *12*(2), 141-165. https://www.rand.org/pubs/reprints/RP223.html

Arquilla, J., & Ronfeldt, D. (1997). *In Athena's Camp: Preparing for Conflict in the Information Age*. https://www.rand.org/pubs/monograph_reports/MR880.html

Arvidsson, M. (2018). Targeting, Gender, and International Posthumanitarian Law and Practice: Framing The Question of the Human in International Humanitarian Law. *Australian Feminist Law Journal*, *44*(1), 9–28. https://doi.org/10.1080/13200968.2018.1465331

Augustine, A. (2014). *The City of God: Volume I*. Project Gutenberg. https://www.gutenberg.org/files/45304/45304-h/45304-h.htm

Augustine, E. (2003). *Letters 1-99: v. 1* (J. E. Rotelle, Ed.; R. J. Teske, Trans.). New City Press.

Baker, D.-P., & Roberts, D. (2007). Extending Just War Theory: The Jus ad Bellum and the Capabilities Approach to Armed Conflict. *Scientia Militaria - South African Journal of Military Studies*, *35*(2). https://doi.org/10.5787/35-2-36

Barry, C., & Christie, L. (2018). The Moral Equality of Combatants. In *The Oxford Handbook of Ethics of War*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199943418.013.28

Bazargan-Forward, S. (2017). Dignity, Self-Respect, and Bloodless Invasions. In *Who Should Die?: The Ethics of Killing in War*. Oxford University Press. https://oxford.universitypressscholarship.com/view/10.1093/oso/9780190495657.001.0001/oso-9780190495657-chapter-8

Bell, Duncan. *Realism*. Encyclopedia Britannica, 2022, https://www.britannica.com/topic/realism-political-and-social-science.

BBC News. (2020). *$5bn lawsuit for tracking*. BBC News. https://www.bbc.com/news/business-52887340

Blum, G., & Luban, D. (2015). Unsatisfying Wars: Degrees of Risk and the Jus Ex Bello. *Ethics*, *125*(3), 751–780. https://doi.org/10.1086/679558

Booth, R. (2014, June 29). Facebook reveals news feed experiment to control emotions. *The Guardian*. https://www.theguardian.com/technology/2014/jun/29/facebook-users-emotions-news-feeds

Bostrom, N. (2011). *Human Enhancement* (J. Savulescu, Ed.; Reprint edition). Oxford University Press, U.S.A.

Braidotti, R. (2013). *The posthuman*. Polity Press.

Britannica, T. E. (2018). *Ought Implies can*. Britannica, T. Editors of Encyclopaedia (2018, May 4). ought implies can. Encyclopedia Britannica. https://www.britannica.com/topic/ought-implies-can

Broad, W. J., Markoff, J., & Sanger, D. E. (2011, January 15). Israeli Test on Worm Called Crucial in Iran Nuclear Delay. *The New York Times*. https://www.nytimes.com/2011/01/16/world/middleeast/16stuxnet.html

Brook, C. (2011). *Byzantine Hades*. https://threatpost.com/byzantine-hades/91837/

Carr, J. (2010). *Did The Stuxnet Worm Kill India's INSAT-4B Satellite?* Forbes. https://www.forbes.com/sites/firewall/2010/09/29/did-the-stuxnet-worm-kill-indias-insat-4b-satellite/

Carson, A., & Yarhi-Milo, K. (2017). Covert Communication: The Intelligibility and Credibility of Signaling in Secret. *Security Studies*, *26*(1), 124–156. https://doi.org/10.1080/09636412.2017.1243921

Catholc Church. (2019). *Catechism of the Catholic Church*. Libreria Editrice Vaticana. https://www.usccb.org/sites/default/files/flipbooks/catechism/VIII/

CEIC. (2020). *Syria Nominal GDP, 1980 – 2022*. CEIC. https://www.ceicdata.com/en/indicator/syria/nominal-gdp

Charusheela, S. (2009). Social analysis and the capabilities approach: a limit to Martha Nussbaum's universalist ethics. *Cambridge Journal of Economics*, *33*(6), 1135–1152. https://doi.org/10.1093/cje/ben027

Chen, T. M. (2010). Stuxnet, the real start of cyber warfare? [Editor's Note]. *IEEE Network*, *24*(6), 2–3. https://doi.org/10.1109/MNET.2010.5634434

Clark, A. (2004). *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford University Press.

Clark, L. (2021). *NHS GP data grab*. Retrieved March 2, 2022, from https://www.theregister.com/2021/06/01/rcgp_urges_greater_communication/

Clarke, R. A., & Knake, R. (2012). *Cyber War: The Next Threat to National Security and What to Do About It* (Reprint edition). Ecco.

Clausewitz, C. von. (1976). *On War* (M. E. Howard & P. Paret, Trans.). Princeton University Press.

Coates, A. J. (2016). *The Ethics of War / A.J. Coates.* (2nd ed.). University Press.

Colonomos, A. (2017). *Proportionality in Warfare as a Political Norm*. Oxford University Press. https://www.oxfordscholarship.com/view/10.1093/oso/9780198796176.001.0001/oso-9780198796176-chapter-10

Correlates of War. (2022). *History of Correlates of War*. Correlates of War. https://correlatesofwar.org/history

Crisp, R. (2021). Well-Being. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2021). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2021/entries/well-being/

Cutting Sword of Justice. (2012). *Shamoon*. Pastebin.Com. https://pastebin.com/HqAgaQRj

Das, K. (2018, October 19). Turkish Citizens Flock to Bitcoin as Lira Plummets. *Block Telegraph*. https://blocktelegraph.io/turkish-citizens-flock-bitcoin-as-lira-plummets/

Davidovic, J. (2016). Should the Changing Character of War Affect Our Theories of War? *Ethical Theory and Moral Practice*, *19*(3), 603–618. https://doi.org/10.1007/s10677-015-9653-x

DiFrancesco, B. (2019). *Return Of The DAO*. Build Blockchain Tech. https://www.buildblockchain.tech/newsletter/issues/no-73-return-of-the-dao

Dixit, Tanmay, et al. “Why and How to Apply Weber’s Law to Coevolution and Mimicry.” *Evolution*, vol. 75, no. 8, 2021, pp. 1906–19, https://doi.org/10.1111/evo.14290.

Duff, R. A. (2009). Philosophy and ‘The Life of the Law.’ *Journal of Applied Philosophy*, *26*(3), 245–258. https://doi.org/10.1111/j.1468-5930.2009.00449.x

Dvorsky, G. (2014). *How to Comprehend Incomprehensibly Large Numbers*. https://gizmodo.com/how-to-comprehend-incomprehensibly-large-numbers-1531604757

Elster, J. (1989). *Solomonic Judgements: Studies in the Limitation of Rationality*. Cambridge University Press.

Eng, E., & Caselden, D. (2015). *Operation Clandestine Wolf*. https://www.mandiant.com/resources/operation-clandestine-wolf-adobe-flash-zero-day

Epifanova, A. (2020). *Deciphering Russia’s “Sovereign Internet Law” | DGAP*. https://dgap.org/en/research/publications/deciphering-russias-sovereign-internet-law

Evans, G. (2018, June 11). The problem with GPS in the modern military. *Army Technology*. https://www.army-technology.com/features/the-problem-with-gps/

Faesen, L., & Roggeman, A. (2019). *Advancing Cyberstability*. Global Commission on the Stability of Cyberspace. https://cyberstability.org/wp-content/uploads/2019/11/GCSC-Fact-Sheet.pdf

Falliere, N., O Murchu, L., & Olson, E. (2010). *W32.Stuxnet Dossier*. Symantec. https://css.csail.mit.edu/6.858/2014/readings/stuxnet.pdf

Feinberg, J. (1990). *The Moral Limits of the Criminal Law Volume 4: Harmless Wrongdoing*. Oxford University Press. https://www.oxfordscholarship.com/view/10.1093/0195064704.001.0001/acprof-9780195064704

Ferzan, K. K. (2005). Justifying Self-Defense. *Law and Philosophy*, *24*(6), 711–749. https://doi.org/10.1007/s10982-005-0833-z

Finkle, J. (2013). *Researchers say Stuxnet was deployed against Iran in 2007 | Reuters*. https://www.reuters.com/article/idUSL1N0BQ4O720130226

Fischerkeller, M. P., & Harknett, R. J. (2017). Deterrence is Not a Credible Strategy for Cyberspace. *Orbis*, *61*(3), 381–393. https://doi.org/10.1016/j.orbis.2017.05.003

Fisher, D. (2018). *'We Got to Be Cool About This': An Oral History of the L0pht, Part 1*. Decipher. https://duo.com/decipher/an-oral-history-of-the-l0pht

Fiveash, K. (2014). *Ill communication delays NHS England's GP data grab for six months*. https://www.theregister.com/2014/02/19/nhs_england_delays_gp_data_grab_for_six_months_after_uproar/

Fletcher, G. P. (1973). Proportionality and the Psychotic Aggressor: A Vignette in Comparative Criminal Theory*. *Israel Law Review*, *8*(3), 367–390. https://doi.org/10.1017/S002122370000426X

Floridi, L. (2016). *The Fourth Revolution: How the Infosphere is Reshaping Human Reality* (Reprint edition). OUP Oxford.

Foot, P. (1967). The Problem of Abortion and the Doctrine of the Double Effect. *Oxford Review*, *5*, 5-15.

French, S. E. (2018). Distinction and Civilian Immunity. In L. May (Ed.), *The Cambridge Handbook of the Just War* (pp. 152–166). Cambridge University Press. https://doi.org/10.1017/9781316591307.010

Frowe, H. (2014). Non-Combatant Liability in War*. In *How We Fight*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199673438.003.0010

Frowe, H. (2015). *Can Reductive Individualists Allow Defense Against Political Aggression?* Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199669530.001.0001/acprof-9780199669530-chapter-8

Fruhlinger, J. (2017). *What is Stuxnet, who created it and how does it work?* CSO Online. https://www.csoonline.com/article/3218104/what-is-stuxnet-who-created-it-and-how-does-it-work.html

Gady, F.-S. (2015). *New Snowden Documents Reveal Chinese Behind F-35 Hack*. https://thediplomat.com/2015/01/new-snowden-documents-reveal-chinese-behind-f-35-hack/

Galliott, J. (Ed.). (2019). *Force Short of War in Modern Conflict: Jus Ad Vim*. Edinburgh University Press. https://www.jstor.org/stable/10.3366/j.ctvggx3k3

Geers, K. (2009). Changing Nature of Warfare. *CCDCOE*, 12.

Gengler, B. (1999). Super-hacker Kevin Mitnick takes a plea. *Computer Fraud & Security*, *1999*(5), 6. https://doi.org/10.1016/S1361-3723(99)90141-0

Golshan, T. (2016, September 26). *First presidential debate transcript*. Vox. https://www.vox.com/2016/9/26/13065174/first-presidential-debate-live-transcript-clinton-trump

Government of Estonia. (2022). *What is e-Residency | How to Start an EU Company Online*. E-Residency. https://www.e-resident.gov.ee/

Greenemeier, L. (2016). *GPS and the World's First "Space War."* Scientific American. https://www.scientificamerican.com/article/gps-and-the-world-s-first-space-war/

Gregg, M. (2017). *"Cyber Sabotage" Could Be the Next Big Crime Wave | HuffPost*. https://www.huffpost.com/entry/cyber-sabotage-could-be-the-next-big-crime-wave_b_59482138e4b04d8767077ad4

Hájek, A., & Rabinowicz, W. (2021). Degrees of commensurability and the repugnant conclusion. *Noûs*, *n/a*(n/a). https://doi.org/10.1111/nous.12388

Haque, A. A. (2017). *A Theory of Jus in Bello Proportionality*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/oso/9780198796176.001.0001/oso-9780198796176-chapter-9

Haraway, D. (1990). *Simians, Cyborgs, and Women: The Reinvention of Nature*. Routledge. https://doi.org/10.4324/9780203873106

Harper, D. (n.d.). *Etymology of just*. Online Etymology Dictionary. Retrieved February 16, 2022, from https://www.etymonline.com/word/just

Haslam, J. (2002). *No Virtue Like Necessity: Realist Thought in International Relations Since Machiavelli*. Yale University Press.

Hathaway, O., & Shapiro, S. (2017). *The Internationalists: And Their Plan to Outlaw War* (01 edition). Allen Lane.

Healey, J. (2011). The Spectrum of National Responsibility for Cyberattacks. *The Brown Journal of World Affairs*, *18*(1), 57–70. JSTOR. https://www.jstor.org/stable/24590776

Healey, J. (2012). Beyond Attribution: Seeking National Responsibility in Cyberspace. *Atlantic Council*. https://www.atlanticcouncil.org/in-depth-research-reports/issue-brief/beyond-attribution-seeking-national-responsibility-in-cyberspace/

Healey, J. (Ed.). (2013). *A fierce domain: conflict in cyberspace, 1986 to 2012*. Cyber Conflict Studies Association.

Hemmingway, E. (1946). Introduction. In *Treasury for the Free World*. Arco Publishing Company.

Henderson, S., Miller, S., Perez, D., Siedlarz, M., Wilson, B., & Read, B. (2018). *Chinese Espionage Group TEMP.Periscope Targets Cambodia Ahead of July 2018 Elections and Reveals Broad Operations Globally | Mandiant*. https://www.mandiant.com/resources/chinese-espionage-group-targets-cambodia-ahead-of-elections

Hildebrandt, M. (2016). *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology*. Edward Elgar Publishing Ltd.

Hobbes, T. (1651). *Philosophicall rudiments concerning government and society*. Printed by J.G. for R. Royston ...

Hobbes, T. (1968). *Leviathan / Thomas Hobbes ; edited with an introduction by C.B. Macpherson.* Penguin.

Holder, E. (2014). *Attorney General Eric Holder – Statement on Chinese Hacking*. Genius. https://genius.com/Attorney-general-eric-holder-statement-on-chinese-hacking-annotated

Hurka, T. (2005). *Proportionality in the Morality of War*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:osobl/9780199743094.001.0001/acprof-9780199743094-chapter-15

Hurka, T. (2008). Proportionality and Necessity. In L. May & E. Crookston (Eds.), *War: Essays in Political Philosophy*. Cambridge University Press.

ICRC. (2022a). *Customary IHL - 84. The Protection of Civilians and Civilian Objects from the Effects of Incendiary Weapons*. ICRC. https://ihl-databases.icrc.org/customary-ihl/eng/docs/v2_cou_il_rule84

ICRC. (2022b). *Customary IHL - Practice Relating to Rule 107. Spies*. Customary IHL. https://ihl-databases.icrc.org/customary-ihl/eng/docs/v2_rul_rule107_sectionb

Ilves, T. H. (2007). *Address by the President of Estonia (62 nd Session of the United Nations General Assembly*. United Nations. http://www.un.org/webcast/ga/62/2007/pdfs/estonia-eng.pdf

Insikt. (2017, May 17). Recorded Future Research Concludes Chinese Ministry of State Security Behind APT3. *Recorded Future*. https://www.recordedfuture.com/chinese-mss-behind-apt3/

Izard, V., & Dehaene, S. (2008). Calibrating the mental number line. *Cognition*, *106*(3), 1221–1247. https://doi.org/10.1016/j.cognition.2007.06.004

Jaggar, A. M. (2006). Reasoning About Well-Being: Nussbaum's Methods of Justifying the Capabilities*. *Journal of Political Philosophy*, *14*(3), 301–322. https://doi.org/10.1111/j.1467-9760.2006.00253.x

Javers, E. (2022, March 1). *Ukraine asked the internet's governing body to remove Russian sites*. CNBC. https://www.cnbc.com/2022/03/01/ukraine-asked-icann-to-revoke-russian-domains-shut-dns-servers.html

Jenkins, R. (2016). Cyberwarfare as Ideal War. In *Binary Bullets: The Ethics of Cyberwarfare*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780190221072.001.0001/acprof-9780190221072-chapter-6

Jenkins, R., Robillard, M., & Strawser, B. J. (2017). *Who Should Die?: The Ethics of Killing in War*. Oxford University Press.

Johnson, A. L. (n.d.). *Endpoint Protection - Symantec Enterprise*. Retrieved March 3, 2022, from https://community.broadcom.com/symantecenterprise/communities/community-home/librarydocuments/viewdocument?DocumentKey=281521ea-2d18-4bf9-9e88-8b1dc41cfdb6&CommunityKey=1ecf5f55-9545-44d6-b0f4-4e4a7f5f5e68&tab=librarydocuments

Jones, A. H., Dizon, Z. B., & October, T. W. (2018). Investigation of Public Perception of Brain Death Using the Internet. *Chest*, *154*(2), 286–292. https://doi.org/10.1016/j.chest.2018.01.021

Kahn, L. (2018). *Liability to Deadly Force in War*. Oxford University Press. https://oxford.universitypressscholarship.com/view/10.1093/oso/9780190495657.001.0001/oso-9780190495657-chapter-2

Kaldor, M. (2007). *New and Old Wars*. Stanford University Press.

Kharpal, A. (2021, March 4). *China has given away millions in its digital yuan trials. This is how it works*. CNBC. https://www.cnbc.com/2021/03/05/chinas-digital-yuan-what-is-it-and-how-does-it-work.html

Killings, D. (1996). *The Anglo Saxon Chronicle*. The Anglo Saxon Chronicle. https://www.fulltextarchive.com/pdfs/The-Anglo-Saxon-Chronicle.pdf

Kirk, R. (2021). Zombies. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2021). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2021/entries/zombies/

Korab-Karpowicz, W. J. (2018). Political Realism in International Relations. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2018). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2018/entries/realism-intl-relations/

Kuerbis, B. (2018). Research on public attribution of state-sponsored attacks. *Internet Governance Project*.

https://www.internetgovernance.org/2018/10/08/research-on-public-attribution-of-state-sponsored-attacks/

Kutz, C. (2005). The Difference Uniforms Make: Collective Violence in Criminal Law and War. *Philosophy & Public Affairs*, *33*(2), 148–180. https://doi.org/10.1111/j.1088-4963.2005.00028.x

Landy, D., Charlesworth, A., & Ottmar, E. (2014). Cutting In Line: Discontinuities in the Use of Large Numbers by Adults. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *36*(36). https://escholarship.org/uc/item/7tb2h41q

Lange, J. (2016). *Clinton appears to suggest rigging the Palestine election*. The Week. https://theweek.com/speedreads/658320/unearthed-2006-audio-clinton-appears-suggest-rigging-palestine-election

Langner, R. (2013). Stuxnet's Secret Twin. *Foreign Policy*. https://foreignpolicy.com/2013/11/19/stuxnets-secret-twin/

Lazar, S. (2014). National Defence, Self-Defence, and the Problem of Political Aggression. In C. Fabre & S. Lazar (Eds.), *The Morality of Defensive War* (pp. 11–39). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199682836.003.0002

Lazar, S. (2015). *Sparing Civilians*. Oxford University Press.

Lazar, S. (2017). Just War Theory: Revisionists Versus Traditionalists. *Annual Review of Political Science*, *20*(1), 37–54. https://doi.org/10.1146/annurev-polisci-060314-112706

Lazar, S. (2020). War. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2020). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2020/entries/war/

Leavitt, H. J., & Whisler, T. L. (1958). Management in the 1980's. *Harvard Business Review*. https://hbr.org/1958/11/management-in-the-1980s

Lucas, G. (2017). *Ethics and Cyber Warfare: The Quest for Responsible Security in the Age of Digital Warfare* (1 edition). OUP USA.

Mandela, N. (1964). *I am prepared to die*. Nelson Mandela Foundation. http://db.nelsonmandela.org/speeches/pub_view.asp?pg=item&ItemID=NMS010

Mandiant. (2021). *Advanced Persistent Threat Groups | Mandiant*. https://www.mandiant.com/resources/apt-groups

Mantovani, M. (2017). Francisco de Vitoria on the "Just War": Brief Notes and Remarks. In J. M. Beneyto & J. Corti Varela (Eds.), *At the Origins of Modernity:*

Francisco de Vitoria and the Discovery of International Law (pp. 119–139). Springer International Publishing. https://doi.org/10.1007/978-3-319-62998-8_7

Markoff, J. (1994). Cyberspace's Most Wanted: Hacker Eludes F.B.I. Pursuit. *The New York Times*. https://www.nytimes.com/1994/07/04/us/cyberspace-s-most-wanted-hacker-eludes-fbi-pursuit.html

Mattox, J. M. (2009). *St. Augustine and the Theory of Just War*. Bloomsbury Publishing Plc.

Mattox, J. M. (2011). Augustine: Political and Social Philosophy. In *Internet Encyclopedia of Philosophy*. https://iep.utm.edu/aug-poso/

Mattox, J. M. (2018). *The Just War Tradition in Late Antiquity and the Middle Ages*. The Cambridge Handbook of the Just War. /core/books/cambridge-handbook-of-the-just-war/just-war-tradition-in-late-antiquity-and-the-middle-ages/53A0B5B9EC661AF938FBF552B4C9BA95

May, L. (2018). Introduction. In L. May (Ed.), *The Cambridge Handbook of the Just War* (pp. 1–10). Cambridge University Press. https://doi.org/10.1017/9781316591307.002

McAfee. (2019). *What Is Stuxnet? | McAfee*. https://www.mcafee.com/enterprise/en-us/security-awareness/ransomware/what-is-stuxnet.html

McElwee, J. J. (2020). *Catholic activists praise pope's move away from just war theory*. National Catholic Reporter. https://www.ncronline.org/news/justice/catholic-pacifists-praise-popes-move-away-just-war-theory

McMahan, J. (2005). The Basis of Moral Liability to Defensive Killing. *Philosophical Issues*, *15*(1), 386–405. https://doi.org/10.1111/j.1533-6077.2005.00073.x

McMahan, J. (2009). *Killing in war / Jeff McMahan.* University Press.

McMahan, J. (2015). Proportionality and Time. *Ethics*, *125*(3), 696–719. https://doi.org/10.1086/679557

Mellow, D. (2006). Counterfactuals and the Proportionality Criterion. *Ethics & International Affairs; New York*, *20*(4), 439-454,541. https://doi.org/http://dx.doi.org.ezproxy.nottingham.ac.uk/10.1111/j.1747-7093.2006.00044.x

Merriam-Webster. (n.d.). espionage, n. In *Merriam-Webster.com dictionary*. Retrieved March 9, 2022, from https://www.merriam-webster.com/dictionary/espionage

Miller, Z. (2014, July 29). *Israeli Ambassador: Here's What "Proportionality" In War Really Means | TIME*. https://web.archive.org/web/20140729055624/https://time.com/

Mimoso, M. (2014). *Stuxnet's First Five Victims Provided Path to Natanz*. https://threatpost.com/stuxnets-first-five-victims-provided-path-to-natanz/109291/

Mitre Att&ck. (2017). *APT3, Gothic Panda, Pirpi, UPS Team, Buckeye*. https://attack.mitre.org/groups/G0022/

Mollendorf, D. (2008). Jus ex Bello*. *Journal of Political Philosophy*, *16*(2), 123–136. https://doi.org/https://doi.org/10.1111/j.1467-9760.2008.00310.x

Morton, H. (2019, January 23). *Net Neutrality Legislation in States*. https://www.ncsl.org/research/telecommunications-and-information-technology/net-neutrality-legislation-in-states.aspx

Neu, M. (2012). Why McMahan's just wars are only justified and why that matters. *Ethical Perspectives*, *19*(2), 235–255. https://doi.org/10.2143/EP.19.2.2160705

Neu, M. (2013). The tragedy of justified war. *International Relations*, *27*(4), 461–480. https://doi.org/10.1177/0047117813483434

N.H.S. (2022). *Digitisation - Primary Care Support England*. N.H.S. England. https://pcse.england.nhs.uk/services/medical-records/digitisation/

Nussbaum, M. C. (1995). Introduction. In M. C. Nussbaum & J. Glover (Eds.), *Women, Culture, and Development: A Study of Human Capabilities*. https://academic.oup.com/book/12441/chapter/162053526

Nussbaum, M. C. (2000). *Sex and Social Justice*. Oxford University Press.

OED Online. (2022a). advantage, n. In *OED Online*. Oxford University Press. https://www.oed.com/view/Entry/2895

OED Online. (2022b). espionage, n. In *OED Online*. Oxford University Press. https://www.oed.com/view/Entry/64416

OED Online. (2022c). sabotage, n. In *OED Online*. Oxford University Press. https://www.oed.com/view/Entry/169373

OED Online. (2022d). spy, n. In *OED Online*. Oxford University Press. https://www.oed.com/view/Entry/188063

O'Neill, O. (2016). Universalism in ethics. In *Routledge Encyclopedia of Philosophy* (1st ed.). Routledge. https://doi.org/10.4324/9780415249126-L108-1

Orend, B. (2019). *War and Political Theory*. Polity Press.

Otsuka, M. (1994). Killing the Innocent in Self-Defense. *Philosophy & Public Affairs*, *23*(1), 74–94. JSTOR. http://www.jstor.org/stable/2265226

Page, J. (2021). *Cryptocurrency Growth Statistics*. Cryptohead.Io. http://web.archive.org/web/20210615152137/https://cryptohead.io/learn-crypto/cryptocurrency-growth-statistics/

Pagliery, J. (2015). *The inside story of the biggest hack in history*. https://money.cnn.com/2015/08/05/technology/aramco-hack/index.html

Parry, J. (2015). Liability, community, and just conduct in war. *Philosophical Studies*, *172*(12), 3313–3333. https://doi.org/10.1007/s11098-015-0471-8

Pauli, D. (2013). *Stuxnet infected Russian nuclear plant*. ITnews. https://www.itnews.com.au/news/stuxnet-infected-russian-nuclear-plant-363578

Pinker, S. (2014, February 16). *Daniel Kahneman changed the way we think about thinking. But what do other thinkers think of him?* https://www.theguardian.com/science/2014/feb/16/daniel-kahneman-thinking-fast-and-slow-tributes

Pope Francis. (2020, October 3). *Fratelli tutti*. Vatican.Va. https://www.vatican.va/content/francesco/en/encyclicals/documents/papa-francesco_20201003_enciclica-fratelli-tutti.html

Primoratz, I. (Ed.). (2007). *Civilian Immunity in War*. Oxford University Press.

Privitera, J. (2018). Aggregate Relevant Claims in Rescue Cases? *Utilitas*, *30*(2), 228–236. https://doi.org/10.1017/S0953820817000279

Quong, J. (2012). Liability to Defensive Harm. *Philosophy & Public Affairs*, *40*(1), 45–77. https://doi.org/10.1111/j.1088-4963.2012.01217.x

Rasmussen, D. M. (1990). *Universalism Vs. Communitarianism: Contemporary Debates in Ethics*. MIT Press.

Ray, M. (n.d.). *8 Deadliest Wars of the 21st Century*. Encyclopedia Britannica. Retrieved February 24, 2022, from https://www.britannica.com/list/8-deadliest-wars-of-the-21st-century

Reichberg, G. M. (2011b). Suárez on just war. In D. Schwartz (Ed.), *Interpreting Suárez: Critical Essays* (pp. 185–204). Cambridge University Press. https://doi.org/10.1017/CBO9781139018753.009

Reitman, R. (2021). *The Cryptocurrency Surveillance Provision Buried in the Infrastructure Bill is a Disaster for Digital Privacy*. Electronic Frontier Foundation. https://www.eff.org/deeplinks/2021/08/cryptocurrency-surveillance-provision-buried-infrastructure-bill-disaster-digital

Reuters. (2010, November 29). Iran says cyber foes caused centrifuge problems. *Reuters*. https://www.reuters.com/article/idUSLDE6AS1L120101129

Rid, T. (2012). Cyber War Will Not Take Place. *Journal of Strategic Studies*, *35*(1), 5–32. https://doi.org/10.1080/01402390.2011.608939

Rid, T. (2013). *Cyber War Will Not Take Place*. Oxford University Press, Incorporated. http://ebookcentral.proquest.com/lib/nottingham/detail.action?docID=1364050

Ripstein, A. (2006). Beyond the Harm Principle. *Philosophy & Public Affairs*, *34*(3), 215–245. https://doi.org/10.1111/j.1088-4963.2006.00066.x

Robeyns, I. (2016). The Capability Approach. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2016/entries/capability-approach/

Rodin, D. (2002). *War and Self-Defense*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/0199257744.001.0001/acprof-9780199257744-chapter-1

Rodin, D. (2004). War and Self-Defense. *Ethics & International Affairs*, *18*(1), 63–68. https://doi.org/http://dx.doi.org/10.1111/j.1747-7093.2004.tb00451.x

Rodin, D. (2014). *The Myth of National Self-Defence*. Oxford University Press. https://www-oxfordscholarship-com.ezproxy.nottingham.ac.uk/view/10.1093/acprof:oso/9780199682836.001.0001/acprof-9780199682836-chapter-4

Rodin, D. (2015). The War Trap: Dilemmas of jus terminatio. *Ethics*, *125*(3), 674–695. https://doi.org/10.1086/679559

Ryan, C. (2013). Pacifism, Just War, and Self-Defense. *Philosophia*, *41*(4), 977–1005. https://doi.org/10.1007/s11406-013-9493-7

Samani, R., & Beek, C. (2017, April 26). Shamoon Returns, Bigger and Badder. *McAfee Blog*. https://www.mcafee.com/blogs/enterprise/shamoon-returns-bigger-badder/

Sanger, D. E. (2014, May 19). With Spy Charges, U.S. Draws a Line That Few Others Recognize. *The New York Times*. https://www.nytimes.com/2014/05/20/us/us-treads-fine-line-in-fighting-chinese-espionage.html

Savelyev, A. (2016). *Contract Law 2.0: «Smart» Contracts As the Beginning of the End of Classic Contract Law* (SSRN Scholarly Paper ID 2885241). Social Science Research Network. https://papers.ssrn.com/abstract=2885241

Schleifer, T., & Walsh, D. (2016). *Russian cyberintrusions an "act of war" | CNN*. https://www.cnn.com/2016/12/30/politics/mccain-cyber-hearing/index.html

Schmitt, M. N. (1999). *Computer Network Attack and the Use of Force in International Law: Thoughts on a Normative Framework* (SSRN Scholarly Paper ID 1603800). Social Science Research Network. https://papers.ssrn.com/abstract=1603800

Schmitt, M. N. (2013). *Tallinn Manual on the International Law Applicable to Cyber Warfare* (Reprint edition). Cambridge University Press.

Seals, T. (2019). *Stuxnet-Related APTs Form Gossip Girl*. https://threatpost.com/stuxnet-apts-gossip-girl/143595/

Sen, A. (2004). Capabilities, Lists, and Public Reason: Continuing the Conversation. *Feminist Economics*, *10*(3), 77–80. https://doi.org/10.1080/1354570042000315163

Shears, M., Daniela, Schnidrig, D., & Kaspar, L. (2018). *Multistakeholder Approaches to National Cybersecurity Strategy Development*. Global Partners Digital. https://www.gp-digital.org/wp-content/uploads/2018/06/Multistakeholder-Approaches-to-National-Cybersecurity-Strategy-Development.pdf

Sigalos, M. (2021, July 7). *China's war on bitcoin just hit a new level with its latest crypto crackdown*. CNBC. https://www.cnbc.com/2021/07/06/china-cracks-down-on-crypto-related-services-in-ongoing-war-on-bitcoin.html

Simon, W. L., Mitnick, K. D., & Wozniak, S. (2003). *The Art of Deception: Controlling the Human Element of Security* (1st edition). Wiley.

Small, M., & Singer, J. D. (1982). *Resort to arms: international and civil wars, 1816-1980* (2nd ed.). Sage Publications.

Smith, T. (2001). *Hacker jailed for revenge sewage attacks*. https://www.theregister.com/2001/10/31/hacker_jailed_for_revenge_sewage/

SolarWinds. (2021). *Security Advisory*. SolarWinds. https://www.solarwinds.com/sa-overview/securityadvisory

Sreedhar, S. (2008). Defending the Hobbesian Right of Self-Defense. *Political Theory*, *36*(6), 781–802. https://www.jstor.org/stable/20452669

Statista Research Department. (2021). *The Syrian Civil War*. Statista. https://www.statista.com/topics/4216/the-syrian-civil-war/

Statman, D. (2015). Ending War Short of Victory? A Contractarian View of Jus Ex Bello. *Ethics*, *125*(3), 720–750. https://doi.org/10.1086/679561

Steinhoff, U. (2007). *On the Ethics of War and Terrorism*. Oxford University Press. https://oxford-universitypressscholarship-com.ezproxy.nottingham.ac.uk/view/10.1093/acprof:oso/9780199217373.001.0001/acprof-9780199217373

Steinhoff, U. (2009, April 1). *What Is War—And Can a Lone Individual Wage One?* International Journal of Applied Philosophy. https://www.pdcnet.org/pdc/bvdb.nsf/purchase?openform&fp=ijap&id=ijap_2009_0023_0001_0133_0150

Steinhoff, U. (2014). Just Cause and 'Right Intention.' *Journal of Military Ethics*, *13*(1), 32–48. https://doi.org/10.1080/15027570.2014.908647

Suárez, F. (1858). *Opera omnia*. Apud Ludovicus Vivès.

Taddeo, M. (2012). Information Warfare: A Philosophical Perspective. *Philosophy & Technology*, *25*(1), 105–120. https://doi.org/10.1007/s13347-011-0040-9

Tadros, V. (2012). Duty and Liability. *Utilitas*, *24*(2), 259–277. https://doi.org/10.1017/S095382081200012X

thedryes. (2013). *30 Best Spy Films | IMDb*. IMDb. https://www.imdb.com/list/ls053334053/

Thomson, J. J. (1976). *Self-Defense and Rights*. University of Kansas, Department of Philosophy. https://kuscholarworks.ku.edu/handle/1808/12392

Thomson, J. J. (1991). Self-Defense. *Philosophy & Public Affairs*, *20*(4), 283–310. JSTOR. https://www.jstor.org/stable/2265419

Tomlin, P. (2017). On Limited Aggregation. *Philosophy & Public Affairs*, *45*(3), 232–260. https://doi.org/10.1111/papa.12097

Tomlin, P. (2018). Subjective Proportionality. *Ethics*, *129*(2), 254–283. https://doi.org/10.1086/700031

Tornau, C. (2020). Saint Augustine. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2020). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2020/entries/augustine/

Tsagourias, N. (2019). Electoral Cyber Interference, Self-Determination and the Principle of Non-Intervention in Cyberspace. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3438567

UK Government. (2006). *Protocol additional to the Geneva Conventions of 12 August 1949*. GOV.UK. https://www.gov.uk/government/publications/protocol-additional-to-the-geneva-conventions-of-12-august-1949

UK Government. (2021). *Smart contracts | Law Commission*. https://www.lawcom.gov.uk/project/smart-contracts/

UN News. (2021). *Syria: 10 years of war has left at least 350,000 dead*. UN News. https://news.un.org/en/story/2021/09/1101162

United Nations. (1945a). *Charter of the United Nations*. Article 51. https://legal.un.org/repertory/art53.shtml

United Nations. (1945b). *Charter of the United Nations*. Article 2. https://legal.un.org/repertory/art2.shtml

United Nations. (2020). *Operationalizing Cyber Norms: Multi-stakeholder Approaches to Responsible Vulnerabilities Disclosure (January 20, 2020)*. Indico - Accreditation System. https://indico.un.org/event/33261/

U.S. Government. (1995). *#089 | US Gov*. Fugitive Computer Hacker Arrested in North Carolina. https://www.justice.gov/archive/opa/pr/Pre_96/February95/89.txt.html

U.S. Government. (2016a). *How to Protect Your Networks from Ransomware*. https://www.justice.gov/criminal-ccips/file/872771/download

U.S. Government. (2016b). *Shamoon/DistTrack Malware | CISA*. https://www.cisa.gov/uscert/ics/jsar/JSAR-12-241-01B

U.S. Government. (2017a). *Background to "Assessing Russian Activities and Intentions in Recent US Elections."*

U.S. Government. (2017b). *SEC Issues Investigative Report Concluding DAO Tokens, a Digital Asset, Were Securities*. https://www.sec.gov/news/press-release/2017-131

U.S. Government. (2018). *National Cyber Strategy*. U.S. Government. https://trumpwhitehouse.archives.gov/wp-content/uploads/2018/09/National-Cyber-Strategy.pdf

U.S. Legal. (n.d.). *Sabotage Law and Legal Definition*. U.S.. Retrieved March 3, 2022, from https://definitions.uslegal.com/s/sabotage/

Valeriano, B., Jensen, B., & Maness, R. C. (2018). *Cyber Strategy: The Evolving Character of Power and Coercion*. Oxford University Press.

Vigliarolo, B. (2017). *Stuxnet: The smart person's guide*. TechRepublic. https://www.techrepublic.com/article/stuxnet-the-smart-persons-guide/

Vitoria, F. de. (1964). *De Indis Et de Iure Belli Relectiones*. Oceana.

Voorhoeve, A. (2014). How Should We Aggregate Competing Claims? *Ethics*, *125*(1), 64–87. JSTOR. https://doi.org/10.1086/677022

Voorhoeve, A. (2018). Balancing small against large burdens. *Behavioural Public Policy*, *2*(1), 125–142. https://doi.org/10.1017/bpp.2017.4

Walzer, M. (2006a). *Just and Unjust Wars: A Moral Argument with Historical Illustrations* (4 edition). Basic Books.

Walzer, M. (2006b). Response to McMahan's Paper. *Philosophia*, *34*, 43–45. https://doi.org/10.1007/s11406-006-9008-x

Watkin, K. (2005). Assessing Proportionality: Moral Complexity and Legal Rules. *Yearbook of International Humanitarian Law*, *8*, 3–53. https://doi.org/10.1017/S1389135905000036

Weinberger, S. (2007). How Israel Spoofed Syria's Air Defense System. *Wired*. https://www.wired.com/2007/10/how-israel-spoo/

Wikipedia. (2022a). Abstraction. In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Abstraction&oldid=1069434421

Wikipedia. (2022b). Information technology. In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Information_technology&oldid=1072737642

Woodard, C. (2013). Classifying theories of welfare. *Philosophical Studies*, *165*(3), 787–803. https://doi.org/10.1007/s11098-012-9978-4

Zetter, K. (2014). An Unprecedented Look at Stuxnet, the World's First Digital Weapon. *Wired*. https://www.wired.com/2014/11/countdown-to-zero-day-stuxnet/