Moral Cyber Weapons[1]

Dorothy E. Denning
Naval Postgraduate School
dedennin@nps.edu

Bradley J. Strawser
Naval Postgraduate School
bjstraws@nps.edu

Abstract

This paper examines the morality of cyber weapons, offering conditions under which they are not only ethical under just war theory, but morally preferred over their kinetic counterparts. When these conditions are satisfied, states not only have the option of using cyber weapons, but could even acquire a moral duty to do so over other forms of warfare. In particular, we show that states are morally obliged to use cyber weapons instead of kinetic weapons when they can be deployed for a purpose already deemed just under the law of armed conflict and without any significant loss of capability. The reason behind this moral obligation is that cyber weapons can reduce both the risk to one's own (putatively just) military and the harm to one's adversary and non-combatants. The paper discusses this obligation, using examples to illustrate cases where it does or does not apply. It also addresses several objections that have been raised about the use of cyber weapons, showing that they fail to fully counter the obligation to use cyber weapons derived from their reduction of risk and harm properties.

[1] Approved for public release; distribution is unlimited. The views expressed in this document are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

Moral Cyber Weapons

Introduction

The formation of military cyber forces in the United States, China, and other nations has stimulated considerable interest in topics relating to the deployment of cyber weapons in state-level conflicts. One area of particular interest, and the topic of this paper, concerns the ethics of using cyber weapons. So far, most scholarly attention has focused primarily on whether cyber-attacks are legally permissible under the international law of armed conflict (LOAC) (Denning 2008; DoD 1999; Owens et al. 2009; Schmitt 1999, 2010; Wingfield 2008, 2010). LOAC derives from the just war theory tradition and consists of two primary divisions: *jus ad bellum*, the ethical justification for going to war, and *jus in bello*, the moral principles governing conduct within war. Both are concerned with state use of force, particularly armed forces, but the former specifies *when* that force may be applied, whereas the latter specifies ground rules for *how* it should be applied as part of the prosecution of a justified war. Together, they enshrine widely accepted ethical principles that are intended to promote peace and minimize the adverse effects of war on the world; the just war convention can be understood as a restraining influence on the moral horrors of war. However, these just war theory principles governing LOAC come from an era that predates cyberspace, leaving their applicability to cyber weapons a question open to interpretation. Some scholars have argued that cyberwarfare is so divergent from traditional forms of warfare that the principles of just war theory simply do not apply in any straightforward manner to cyber weapons (Dipert 2010). We disagree. We argue that, at least with some kinds of cyber weapons, not only can they adhere to the principles of just war theory but that a positive duty to employ them can arise, at least in certain contexts.

The main principles of *jus ad bellum* are codified in the Charter of the United Nations, which specifies the conditions under which member states may apply force against other states. The most relevant parts of the Charter include Article 2(4), which prohibits states from using force against other states during peacetime; Article 39, which gives the U.N. Security Council responsibility for responding to threats and acts of aggression; and Article 51, which gives states a right to self-defense. With respect to cyber weapons, the primary question has been whether cyber-attacks constitute the use of force and, therefore, fall under the provisions of the above articles. Two general approaches to force analysis have been proposed: equivalent effects analysis (DoD 1999) and the Schmitt analysis (Schmitt 1999, 2010). Equivalent effects analysis considers a cyber-attack to be a use of force if its effects are equivalent to those of an armed attack, while the Schmitt analysis uses a broad set of criteria to distinguish the application of armed force from permissible actions such as trade sanctions. Both can be difficult to apply to cyber-attacks, leaving considerable uncertainty as to whether a particular cyber-

attack constitutes an illegal use of force under Article 2(4) or even whether such attacks could even be properly understood as acts of war (Rid 2011).[2]

Jus in bello principles are concerned with whether operations conducted during a state of war follow the moral principles of necessity, distinction, proportionality, neutrality, perfidy, discrimination, and superfluous injury. Applying these principles to cyber operations has been less problematic, and many believe that cyber weapons could be employed ethically in the context of an otherwise just war (DoD 1999).

Rather than considering the question of whether cyber weapons can be used ethically under LOAC, this paper goes further and argues that, under certain conditions, their use can actually become morally obligatory. When these conditions are satisfied, states not only have the morally permissible option of using cyber weapons, but a moral duty to do so. In particular, we show that states are morally obliged to use cyber weapons *in place of* kinetic weapons for a just attack whenever doing so does not result in a significant loss of capability. The reason for this moral obligation is that cyber weapons reduce both the risk to one's own (putatively just) military and the harm to one's adversary and non-combatants. Overall, cyber weapons are more humane, less destructive, and less risky than kinetic weapons for achieving certain military effects.

The scope of our ethical analysis in this paper is highly constrained. It does not address the larger question of whether an arbitrary cyber-attack is ever permissible or whether such an attack should even constitute a use of force.[3] Instead, we restrict attention to cyber-attacks that are viable alternatives to kinetic attacks that have already been determined to be just under LOAC, given certain criteria. Arguably, if the use of certain kinetic weapons in a specific context is deemed morally permissible, then the use of cyber weapons to achieve all or some of the same effects should likewise be morally permissible. However, we go further than simply saying that deploying cyber weapons in that context is a morally acceptable alternative. We claim that there is a moral *duty* to use the cyber weapons under such circumstances. This moral obligation arises because cyber weapons can, in some cases, incur less risk and result in less harm than their physical counterparts, while still meeting the same level of mission capability of said physical weapons.

We will first discuss the general claim. Then we will discuss some objections that have been raised to the general idea of using cyber-attacks for military objectives that would function as objections to our argument for the moral obligation to use cyber-weapons. Ultimately, we find that the objections, while substantial, fail to fully counter

[2] Rid (2011) has recently given a philosophical argument (rather than a legal argument) that the use of cyber weapons cannot and will not constitute war. He makes this case based on an analysis of the definition of war such that any given attack must be a lethal, instrumental, and political act in order to constitute war. And he finds cyber-attacks, in isolation, would not constitute all three criteria. We set Rid's analysis aside for the purposes of this paper.

[3] See Dipert (2010: 393) for an argument that the preemptive use of cyber weapons would likely be morally preferable over a similar preemptive kinetic attack, and possibly even morally permissible, precisely for the kinds of advantages cyber weapons have that we rely on in this paper (i.e. their general non-lethality and lesser degree of destruction). On the use of force question, see Rid (2011) noted above.

the obligation to use cyber weapons derived from their reduction of risk and harm properties. Our overall approach follows that used by one of us, Strawser, to argue for a moral duty to employ unmanned aerial vehicles (UAVs) (Strawser 2010). This paper draws on Strawser's theory and approach in that work. However, whereas Strawser previously focused exclusively on risk, we here also consider harm.

Cyber Weapons as Ethically Obligatory

Strawser argues for the use of UAVs from the principal of unnecessary risk (PUR), which he formulates as follows:

> PUR: If X gives Y an order to accomplish good goal G, then X has an obligation, other things being equal, to choose a means to accomplish G that does not violate the demands of justice, make the world worse, or expose Y to potentially lethal risk unless incurring such risk aids in the accomplishment of G in some way that cannot be gained via less risky means.

An important aspect of this formulation is that Y is being ordered to accomplish goal G, which itself is a good and fully justified goal worthy of pursuit. The principle does not prohibit Y from choosing a risky approach on his or her own accord. Rather, it only prohibits someone from ordering Y to use an approach than carries with it avoidable risks.

Taking this principle to be uncontroversial,[4] Strawser applies it to the military use of UAVs with the following operating principle (OP):

> OP: For any just action taken by a given military, if it is possible for the military to use UAV platforms in place of inhabited aerial vehicles without a significant loss of capability, then that military has an ethical obligation to do so.

Strawser argues for OP on the grounds that because UAVs do not risk the lives of their remote pilots, militaries are obliged to use them in circumstances where they do not result in a loss of capability to conduct an operation that is otherwise deemed just in accordance with the principles of *jus ad bellum* and *jus in bello*. In short, in the context of a fully justified war effort, it would be wrong for a military commander to order a manned aircraft operation when the same result can be achieved with an unmanned one and doing so does not in any way worsen the warrior's ability to fight justly. This is because to do otherwise would place an unnecessary risk on the warfighter. Strawser also observers that because unmanned missions are generally less costly than manned ones, this could provide further grounds for a moral obligation for militaries to use them, as money is then freed up for more worthy social goals. However, because PUR provides a less contingent

[4] Although we cannot pursue it here, not all have found even this modest principle uncontroversial. Uwe Steinhoff (forthcoming) argues that it is, indeed, controversial and that acceptance of it could involve several normative problems. Steinhoff's objections to the PUR deserve response, but such a discussion lies outside of the scope of this paper.

and ultimately more normatively compelling reason for using UAVs, he uses it for the moral obligation's derivation.

Turning now to the use of cyber weapons, we first observe that, like UAV strikes, cyber-attacks can be launched and conducted remotely, making them less risky to military personnel than when engaging in kinetic strikes. Thus, the principle of PUR applies to cyber-attacks for much the same reason as it would for UAVs. However, certain kinds of cyber-attacks potentially have a further moral advantage in that a given military objective may be achievable without causing any loss of life or physical damage to the adversary or innocent third-parties and noncombatants. That is, cyber weapons could cause considerably less harm than the kinetic weapons they replace, while still accomplishing a justified military objective equally as effectively. Thus, there are two morally compelling reasons to use cyber weapons in place of physical weapons where possible: they can reduce both the risk to one's own (presumably just) military forces and they can reduce the harm incurred to the adversary and others.

As with UAVs, we formulate the moral obligation to use cyber weapons first in terms of a general principle, which now factors in both risk and harm. Calling it the Principle of Unnecessary Risk and Harm (PURH), it states:

> (PURH): If X gives Y an order to accomplish good goal G, then X has an obligation, other things being equal, to choose a means to accomplish G that does not violate the demands of justice, or cause unnecessary harm and incur unnecessary risk, unless incurring such risk or delivering such harm aids in the accomplishment of G in some way that cannot be gained via less risky or less harmful means.

The principle of PURH is nearly identical to PUR, except for the addition of harm as a disvalue to avoid in the ordering of just agents in pursuit of a good goal. The duty to not "cause unnecessary harm" could be seen as derived from the PUR's original demand to not "make the world worse." But the PURH aims to make the avoidance of unnecessary harm an explicit part of the formulation. Two central components of traditional *jus in bello* principles are that of proportionality and military necessity. Combined, these principles are generally taken to mean that just forces ought to use as minimal force and deliver as minimal harm as is militarily necessary for, and proportionate to, accomplishing a given just objective. Military operations in pursuit of a just cause generally result in some harm, at least to the adversary. Yet the just war tradition demands that we avoid unnecessary harm to the extent possible in the prosecution of a just war. If some of the harm of war may be avoidable by using cyber weapons instead of kinetic weapons, it seems that the just war tradition would demand that we so use cyber-weapons, where possible.

A brief discussion on this aspect of the proportionality and necessity constraint on just action within war is worth briefly exploring. Usually the proportionality constraint (*in bello*) works to limit the amount of destructive force permissible to use for a given attack such that the predicted damage done is proportionate to the relative importance of

given objective. That is, that the damage done is "worth" the given objective and not excessive compared to the military advantage gained by the attack. Here the relative import of a given objective will be tied directly to the good it does towards the just cause. For example, using a nuclear bomb to take out one lone enemy soldier (which presumably would do very little to advance the success of the just cause but would cause tremendous collateral damage) would be grossly disproportionate.

But proportionality is very closely linked to our understanding of the *jus in bello* principle of military necessity. Necessity can be seen as a growth out of the restraint implicated by proportionality. It demands not only that the force used is a proper "fit" to the objective, but also that only the minimum amount of force necessary to accomplish a mission is used for any given objective.

That means that warriors fighting on behalf of a just cause ("just soldiers") will bear *different* duties vis-à-vis the strictures of proportionality and necessity depending on the options available to them. Say a group of just soldiers, W, is engaging a set of enemy (and putatively unjust) soldiers, Z. W has available to them three means (A, B, or C) of attacking Z, each of which would be equally effective at meeting the given mission objectives W seeks on behalf of the just cause. A is a large bomb that would obliterate Z, but will also destroy the building Z occupies, cause massive damage to the surrounding country-side such as burning up agricultural fields, destroy other nearby buildings, (unintentionally) kill some nearby noncombatants, and so on. B is a bomb similar to A, but its blast radius will not extend much beyond Z, although it will destroy the building Z is occupying. C is a weapon which will target the individual members of Z, but will not do any damage to the building they occupy or surrounding area whatsoever. Recall: W has high confidence that A, B, and C are each equally likely to succeed and they have equal access to all three choices.[5]

Under such a scenario, proportionality demands that W use C over A and B if they are to be in alignment with *jus in bello* principles. Were W to use A or B in this case, they would be in violation of *jus in bello* and would not be acting justly in war; they have a moral *obligation* to use C and an obligation to *not* use A and B. But if W found themselves in a scenario whereby they *only* had access to A or B, then the obligation against using A and B would not obtain. In that case, of course, they would be obligated to use B, and obligated to not use A. The point is that proportionality and necessity restrict just actors to use only as much force (resulting in as much harm and risk) as is required to accomplish a given act.[6]

[5] And, further, presume that they do not have a scarcity of resources problem such that they must reserve some particular weapons for future missions, etc.

[6] There are, of course, important parallels here to cases of individual self-defense, where proportionality and necessity rule. In fact, many revisionist accounts of just war theory currently on the rise today contend precisely that the moral rules of warfare should track more closely with the moral reality of individual self-defense and, as such, should impose a much stronger "necessity" clause on any given just military action (see, Rodin 2005, for example). Such a discussion, however, is far outside the scope of this paper.

Like PUR, we consider PURH to be relatively uncontroversial and, within the context of war, it follows from the strictures of proportionality, as just shown.[7] If one need not incur unnecessary risk or harm to carry out a just act, one should not. Applying PURH to cyber weapons leads to the following cyber operating principle (CyberOP):

> CyberOP: For any just action taken by a given military, if it is possible for the military to deploy remote cyber-attacks in place of manned kinetic attacks without a significant loss of capability, then that military has an ethical obligation to do so.

As already noted, there will be less risk associated with the remote deployment of cyber weapons than with manned kinetic operations. If weapons are used sufficiently remotely, there will be little or no risk to the life of military personnel using the weapons. This is the same condition we see obtain with UAVs. Indeed, UAV bomb strikes already presently combine cyber operations (controlling the UAV) with physical operations (dropping the bomb).

In addition, cyber weapons can be less harmful than their physical counterparts. They are generally not lethal and often do not cause any permanent damage to physical infrastructure. Cyber-attacks may even be less damaging than electronic warfare strikes that "fry" electronics such as electro-magnetic pulse (EMP) weapons. A cyber-attack that takes out some service such as telecommunications, power, sensors, or alarms need not cause any permanent damage nor harm anyone. If there is no physical damage, targets are more readily restored to their original state after hostilities have ended. Restoration is faster and costs less. It may just be a matter of restoring bits from backup files, though systems may also have to be patched and security enhanced to avoid future attacks. This can be important for stability and reconstruction operations following war and would lend itself favorably to considerations of *jus post bellum*.[8] Dipert (2009) has argued along similar lines that cyber weapons could even be designed in such a way so that the damage they do is easily reversible, similar to creating an antidote to a real-world contagion. Rowe (2010b) also argues for reversibility and offers four different techniques that could potentially be used to achieve it. Whereas rebuilding critical infrastructure such as telecommunications and electricity power grids can take weeks, months, or even years after being destroyed by *physical* weapons, that same infrastructure could be back up and running within hours or days after a cyber-attack.

As with OP, the operating principle CyberOP presumes that cyber-attacks would be used to take an action that is otherwise deemed just according to the principles of *jus ad bellum* and *jus in bello*, and that using the cyber weapons instead of kinetic ones would not result in any significant loss of capability. The phrase "without a significant loss of capability" is crucial to the formulation of CyberOP, implying that the cyber operations could be sufficiently controlled in such a way that neither risk nor harm would increase in the accomplishment of a given objective. If the deployment of cyber weapons

[7] Again, see Steinhoff (forthcoming) for a contrasting view. Presumably Steinhoff's objections to PUR would carry over to PURH.

[8] See Orend (2000) for a comprehensive case for developing principles of *jus post bellum*.

would incur either greater risks to just military forces or cause more harm to adversary forces or non-combatants, then those weapons would thereby be considered less capable than their physical weapon counterparts. In that case, their use would not be mandatory under CyberOP; indeed, their use would likely be impermissible. It's possible they might yet be considered a better alternative on other moral grounds outside the avoidance of unnecessary risk and harm, but further arguments would have to weigh the tradeoffs involved.

Our claims here, of course, do not rest on mere abstract possibilities and speculation about future kinds of weapons. There are cyber weapons which already exist that could potentially fit the demands of CyberOP. Rattray and Healey (2010) give several examples where cyber operations might be used to support special operations or traditional, kinetic military operations. These include using cyber operations to take out adversary alarms or inject false alarms, or to disrupt telecommunications or command and control networks. Conducting these operations remotely using cyber weapons would be less risky to a just force than sending in military personnel to accomplish the same objectives, and likely less damaging to their targets, making them excellent candidates for application of CyberOp.

The principle of CyberOP, however, would not justify many of the cyber warfare scenarios that have been postulated such as, for example, the one in Clarke and Knake (2010: 64-68) that leads to a nationwide power blackout, airline and subway crashes, pipeline explosions, refinery fires, lethal clouds of chlorine gas, network outages, and more, resulting in thousands of civilian deaths. Kinetic strikes that did all this would almost certainly violate the LOAC, ruling out their cyber equivalents. Even an attack that just took down the Internet would likely violate LOAC, as it would necessitate attacking civilian infrastructure around the world, including infrastructure in neutral countries.[9] If one presumes a blanket principle of non-combatant immunity, it is hard to imagine circumstances where such a strike would be considered just under traditional just war theory.[10] However, in those cases where limited kinetic attacks against the Internet might be justified, say to take out a small number of routers in some country for a limited period of time, surely a cyber-attack that temporarily shut down those same routers would be

[9] For a good discussion on cyber warfare attacking non-combatants and the resulting problems, see Lucas (2011b). Lucas' work in that piece is highly compatible with the claims we make in this paper, although we disagree with Lucas over whether some specific instances of cyber-attacks would be permissible. Lucas writes, "… an act of cyber warfare is permissible if it aims primarily at harming military (rather than civilian) infrastructure, degrades an adversary's ability to undertake highly destructive offensive kinetic operations, harms no civilians and/or destroys little or no civilian infrastructure in the process." And, of course, we further differ from Lucas in contending that in some such instances cyber-attacks would not be merely permissible, but obligatory to use in place of similar kinetic attacks.

[10] Several revisionist just war theorists have recently challenged a blanket principle of non-combatant immunity and have argued that some noncombatants could be liable to harm in war. Jeff McMahan (2009) does this most prominently, but others such as Helen Frowe (2011) have also advanced a rejection of total non-combatant immunity. Note that even on these revisionists accounts, however, an attack against the entire Internet would still fall outside of the bounds of just war practices because there would be very little if any discrimination possible amongst non-combatants. Again see Lucas (2011b) for a discussion on the possibilities for discriminate cyberwarfare.

morally preferred over a kinetic strike that physically ruined the routers and killed the persons operating them.

CyberOP would not even justify many lesser operations that have actually taken place, such as the distributed denial of service (DDoS) attacks against Estonia in 2007 that disrupted access to Estonian websites in protest of the relocation of a Soviet-era war memorial (Tikk et al. 2010; Clarke & Knake 2010: 13-16). One reason is that the nature of the dispute did not justify any military action against Estonia to begin with (Lucas 2011b). Since using kinetic weapons against the target websites would not be permitted, using cyber weapons against those targets could not be justified by CyberOP.

Although cyber-attacks need not cause death or physical destruction, attacks that do are not necessarily ruled out by CyberOP. If the harm is no greater than that caused by the just use of kinetic weapons, then the cyber-attack is still preferred, indeed morally obligatory, if it is less risky. Moreover, a cyber-attack that causes equipment to self-destruct, as in the case of Stuxnet, may still be less life-threatening than physical strikes.

Stuxnet is an interesting and important case. On the one hand, the cyber operation enabled the destruction of centrifuges at Iran's Natanz nuclear enrichment facility (Broad, Markoff & Sanger 2011) without risking the lives of those who did it nor the operators at Natanz. In that sense, it was less risky and less harmful than, say, dropping bombs on Natanz. On the other hand, Stuxnet caused considerable collateral damage that a bomb strike would not have caused. In particular, tens of thousands of other systems got hit with the worm (Falliere et al. 2011). In that regard, Stuxnet was less capable than a kinetic strike, and so not morally obligatory under CyberOP. This does not mean that Stuxnet was immoral or not preferred over a kinetic strike, only that it does not lend itself to direct application of CyberOP. The morality of Stuxnet is more complicated and we return to it below.

Objections to cyber warfare

Although we believe there is a strong case for conducting cyber-attacks in limited circumstances where they have a moral advantage over kinetic attacks, some have argued that militaries should not conduct cyber-attacks at all. If that is true, then, of course, CyberOP would be a vacuous principle, at best. The following reviews some of these objections. We find that each objection, while important, fails to overcome the strong normative force of PURH and the resulting CyberOP.

Objection 1: Cyberspace should not be militarized

Some argue that military operations should not be conducted in cyberspace, as doing so makes cyberspace less attractive and usable to others, turning it into a perpetual battleground. At any given time, militaries might conduct operations that impair normal

activity and harm legitimate use. Everyone who uses the internet would potentially become a target or unwittingly caught in cyber-crossfire.

Our response to this objection is that cyberspace is *already* under constant attack by criminals, protestors, patriotic hackers, cyber jihadists, spies, anarchist groups, and others who pay no heed to legal or ethical constraints on their behavior. By contrast, the militaries of states are or should be concerned with these things, and should conduct themselves under the principles of *jus ad bellum* and *jus in bello*. The cyber military operations we are advocating for in this paper would be so conducted. They are unlikely to even be noticed by most users. Their greatest impact would be felt by legitimate military targets, with less collateral damage than from kinetic strikes. If not and the use of cyber weapons in question was conducted contrary to just war principles, then they would fail to be justified under PUHR to begin with since they would not be a proper case of a good goal G pursued in a manner that does not violate the demands of justice.

A related argument is that military use of cyber weapons runs counter to efforts to make cyberspace more secure and usable. This is because militaries classify their cyber weapons and keep them secret. They will not make their weapons public or report the vulnerabilities they exploit, as doing so would lead to the flaws being repaired, thereby rendering the weapons useless (Rowe 2010a). The net effect of this secrecy is that cyberspace will have unnecessary vulnerabilities that can be exploited not just by militaries but by anyone else discovering them.

Although cyber security is critically important, we do not agree that it is jeopardized by military classification or use of cyber weapons. One reason is that the effects of publicly disclosing vulnerabilities and attack tools are not all positive. While disclosure is likely to enhance security in the long run by removing vulnerabilities, it can also lead to the proliferation of cyber weapons as well as an increase of attacks, as tool developers build on each other's work to create new cyber weapons and criminals take advantage of the lag between disclosure and remediation to launch attacks. Another reason is that if militaries have no vested interest in offensive cyber weapons, they will not allocate resources to cyber weapons development, and hence, will have little information to contribute in the area of cyber vulnerabilities. Moreover, vulnerabilities they do find may be classified anyway in order to protect military systems.

Objection 2: The deployment of cyber weapons will lead to their spread and use

The deployment of cyber weapons typically has the side effect of making those weapons available to the target and possibly third parties. This is because some or even all of the weapon's code may be present on devices hit by the weapons. For example, when a worm spreads to some computer, its code will be present on the infected computer, making it available to the computer's owner and possibly third parties such as anti-virus companies. As another example, computers that have been compromised and placed on botnets often download additional code in response to instructions from their

botnet's command and control facility. All of this code is then available to those with access to the machines.

Although much of the code left behind from cyber weapons will be in the form of executable binaries that are not readily re-purposed, executable code can be reverse-engineered or decompiled into source code, making it more readily available for analysis, reuse, and integration into other code. As a result, the confiscated cyber weapon might be reused or used as a building block for new cyber weapons, perhaps ones that are even more damaging than the original. The new weapons then might be fired back at the source of the original weapon or used to attack other targets. Even if the original tool was used justly, its reuse and offspring might be appropriated for unethical purposes. The net effect can be an increase of damaging cyber-attacks. In addition to spreading covertly as in the above examples, the code for worms, viruses, Trojans, botnets, and other forms of cyber weapons spreads through more overt means. Security researchers, at least many outside government agencies, share information and code (including source code) pertaining to cyber vulnerabilities and tools for exploiting these vulnerabilities. They post it on public websites and sell it through legitimate and black cyber markets. The result has been a proliferation of cyber weapons. The security company Symantec reported that they encountered almost 300 million variants of malicious software in 2010 (Symantec 2011).

Because of these proliferation effects, some argue that militaries should not deploy cyber weapons. There is too much danger that the weapons will get into the wrong hands and be used in harmful ways. Their use will just make the problem of cyber defense worse for everyone. There is an intriguing irony to this argument against the military use of cyber weapons in that it is the inverse of the objection described in the previous subsection, where it was argued that the secrecy of these weapons would limit our ability to learn about and repair vulnerabilities in cyberspace in order to make it more secure. This objection makes the opposite point.

We believe that the proliferation of military-grade cyber weapons is a legitimate and larger concern than their secrecy, especially since such weapons may be more sophisticated than many of the weapons used by other actors in cyberspace. Stuxnet's executable binaries are now out in the public domain, where they have been studied and could be used to develop new weapons. However, militaries can minimize the risk of their cyber weapons falling into adversary and third party hands by precisely targeting them.

Another key way this risk could be mitigated would be for weapons developers to program them to self-destruct after completing their objectives. Although it may not be possible to guarantee complete containment and destruction of military cyber weapons, it may be possible to reach an acceptable level of assurance. If not, then the cyber weapons may not meet the threshold for CyberOP to begin with, as they may be significantly less capable than their kinetic counterparts. We believe that more work needs to be done on developing self-destructing cyber weapons to fight against this proliferation worry. If that

is done, and a cyber weapon's proliferation can be contained, then the obligation to employ such weapon can still apply under CyberOP.

Objection 3: Cyber-attacks are too difficult to control and use effectively

Some argue that cyber-attacks cannot be controlled, and thus could lead to unanticipated and unpredictable harms, including collateral damage. They cite such cyber weapons as viruses and worms, which often spread widely and cause considerable disruption in the process (Rowe 2010a). For example, the Slammer worm shut down ATM machines and emergency 911 systems, caused flight delays, and disabled a safety monitoring system at a nuclear power plant. Even Stuxnet, which limited its primary damage to Iran's nuclear facility, infected tens of thousands of other systems in the process of arriving at and delivering its payload.[11]

We do not accept the premise that cyber weapons cannot be controlled. True, *some* weapons, such as worms that attempt to infect as many devices as possible, seem to be out of control, but we do not anticipate the use of such weapons under CyberOP. That is, any weapons – be they cyber or kinetic – that impose *intentionally* indiscriminate damage in this way would not meet even the most meager adherence to *jus in bello* principles. Rather, we anticipate the use of cyber weapons that are tightly controlled and precisely aimed. Such weapons would have the same or greater level of precision and discrimination as kinetic weapons. Any cyber weapon that could not be controlled would likely fail to meet the principle of CyberOP as it would be less capable than a more controllable and discriminate kinetic weapon. Although kinetic weapons can also cause collateral damage, that damage is relatively limited or, at least, more easily predicted and calculated, compared to cyber weapons that can affect systems all over the Internet.

Rowe (2010a) also argues that cyber weapons are too hard to use effectively. They could be unreliable, as new software systems and weapons often are, with the code failing to work as intended, or an attack failing because assumptions about the target were wrong. Further, the effects of cyber-attacks can be difficult to determine or measure. Even victims of cyber-attacks can have trouble assessing their damages. This is a particularly pernicious problem for the just use of cyber weapons since adherence to *jus in bello* principles of both discrimination and proportionality each require at least some degree of damage predictability for a given attack.

Our position is that cyber weapons whose effects cannot be accurately predicted, controlled, and measured would likely fail to satisfy the conditions for their ethical application under CyberOP. Either they would be less capable than their kinetic counterparts, in which case they would fail to meet the requirements for CyberOP, or else they would have no kinetic equivalents, in which case CyberOP would not apply and

[11] Although in this case much of that infection did little real damage to systems it used on its way to delivering its payload. This is a more complicated question regarding what should constitute damage and how such a calculus should be used in analyzing the moral permissibility of particular attacks. We address this issue separately below in Objection 4.

additional moral reasoning would be needed to determine whether their deployment is morally just. However, if, as Arquilla (1999: 393) argues, the targets of a cyber-attack were strictly limited to strategic military targets, such attacks could very well be morally justified and in-line with CyberOP. We believe this kind of discrimination and control with cyber weapons could be attained.

Objection 4: Cyber-attacks involve using unwilling bystanders as accomplices for attacks

This objection is closely related to Objection 3, but poses its own unique difficulties. Rowe (2010a) claims that military use of cyberspace would produce a kind of unnecessary collateral damage, as militaries would compromise civilian computers in order to place them on botnets or spread worms, or to use them as "stepping stones" or "launch pads" for reaching their targets. We agree that such collateral damage generally should be avoided. However, the objection goes beyond such intentional compromises, as the packets deployed in a cyber-attack could flow through routers and along links owned by private companies and residing in neutral countries. This use of presumably unwilling (and usually unwitting) bystanders (routers) might be considered a violation of the principle of neutrality demanded by *jus in bello*.

Part of this concern arises from the network topology of cyberspace compared with the traditional battle-space of land, sea, or air (Dipert 2010, 2012). When a just force launches and delivers a kinetic weapon to an adversary, they often need not directly involve or traverse various third-party sovereign territories in order to deliver the weapon. Yet cyber-weapons are likely to move through routers in third-party countries on their way to their targets, and the paths they take are difficult to control.[12]

Of course, in the kinetic weapon example, a just force *may* indeed on occasion need to move through a third party's airspace, say, on the way to the target. But if a plane or missile or tank or some such traditional weapon did have to travel through another's territory before arriving at the target, international law would require that they seek permission of the traversed state before doing so. One reason for this is that the traversed state would lose any status of neutrality and, therefore, could be legitimately targeted in a counter attack.

At least in principle, the same permission might be sought for using cyberweapons. However, the situation is considerably more complex, as many sovereign states are likely to be involved in the movement of packets, and packets can flow along different routes and through different countries depending on traffic loads and other dynamic properties of cyberspace. On the other hand, the argument can be made that the movement of packets through third party states does not violate the principle of neutrality. Rather, the situation is the same as for general telecommunications, where belligerents do not need permission to make international phone calls that pass through third party telecommunications switches and links, and the countries providing those

[12] Though this need not always be the case. It is possible for a cyber-weapon to directly attack an adversary system without any mediating system whatsoever. But this will be rare.

switches and links are immune from attack as long as their services are provided impartially to all sides (DoD 1999). Since the routers and links of cyberspace essentially implement the same basic communications relay service, they too should be immune from attack as long as they move the packets of all parties without favoring any side. However, permission would still apply if an attacking state wanted to use the servers of another to launch a cyber-attack or to host files that support the attack.

As we've already made clear, we agree that cyber weapons that produce unnecessary collateral damage should not be used. Indeed, if the kinetic weapons they replace can produce the desired effects but with less collateral damage, then the cyber weapons would be considered less capable and thus non-obligatory under CyberOP. The real difficulty raised by this objection, then, is that such calculations of differing kinds of collateral damage will have to measure not mere quantity of those affected by a just cyber-attack, but the *kind* of harm delivered. It is quite possible that a just cyber-attack that affects a relatively large number of noncombatants as unintended collateral damage, but does so only very minimally (say by very temporarily slightly disrupting their internet access or placing a negligible amount of passive code on their system), could be morally preferable to a just kinetic attack that affected a much smaller number of noncombatants as unintended collateral damage but did so to a much greater degree of harm. If two such approaches were the *only* options available to a just force, it seems at least reasonable that the cyber-attack could be considered the option which produces the least amount of harm, if "least" is meant in terms of severity and not extent. Thereby such a cyber-attack could possibly be justified under PURH and consistent with CyberOP. Such decisions would be difficult, but we do not see any in-principle reason why cyber-attacks should be weighed differently than would two alternative kinetic attacks in similar decisions over weapon choice.

Lucas (2011b) has argued that Stuxnet was justified because of the way its primary damage delivered by the worm was highly discriminate (focusing only on the Iranian centrifuges it was designed to damage) and because an alternative kinetic strike would have done far more physical and, likely, *lethal* harm. He does not consider the other systems infected by Stuxnet to have been harmed. In our view, the non-Iranian systems infected with the Stuxnet worm as part of its delivery should properly be considered collateral damage simply because the owners of those systems could make legitimate complaint that they did not want to have a worm on their system and did not want to have to expend resources removing it and assessing possible damages, all of which can be difficult and time consuming. Further, those infected with Stuxnet on its way to delivery could view it as a violation of their sovereign autonomy over their system in that they unwittingly played the part of accomplice to the attack.

Yet, even if our view regarding the collateral damage of Stuxnet is right, Lucas may still be right that in this case this *kind* of collateral damage is morally preferable to the kind of collateral damage that would have been likely incurred in a kinetic attack on the Iranian nuclear facilities. It that is true, and if a strike designed to impede Iranian nuclear capabilities was otherwise deemed just, then it is possible that Stuxnet could fit

the parameters of CyperOP since it would cause less collateral damage than a comparable kinetic strike.

Again, this conclusion would here be taking "less" collateral damage to mean less severe even if not less extensive. Whether that is the morally correct conclusion for how to best weigh different kinds of collateral damage is a matter for another paper and one for which we remain neutral for the purposes of this paper. We do not here address the difficult and complex ways these different kinds of collateral harms should be weighed against one another when alternative means of accomplishing a just attack are available. But, again, we see no in-principle reason why more widely diffused, but less severe, collateral damage could not be morally preferred over more severe damage inflicted to a smaller set of noncombatants.

Note as well, of course, that in this discussion here of Stuxnet we are not arguing that an attack against the nuclear facilities of Iran was necessarily justified to begin with – kinetically or through cyber-weapons – and do not mean to argue for such a conclusion in this paper. Rather, Stuxnet is a good case to examine in light of CyberOP on the *stipulation* that an attack on Iranian centrifuges was an otherwise just attack. Whether that stipulation is *actually* valid is a debate for another paper.

Objection 5: Cyber-attacks could provoke unanticipated responses and escalate conflicts

It is impossible to predict with certainty how an attacked state might respond to a cyber-attack. Indeed, some countries, including the United States, have said that they would consider all options, including using kinetic weapons against the attacking state. A state might even respond with the nuclear option. Clearly, such actions could escalate a conflict.

Because a cyber-attack might provoke a retaliatory cyber or physical strike far in excess of the original attack, some have argued that cyber-attacks should not be used at all. However, any action can potentially provoke an unanticipated, harmful, and disproportionate response. Even the relatively non-aggressive plan to relocate a war memorial in Estonia provoked not only the DDoS attacks against Estonian websites, but also riots in the streets (Tikk et al. 2010).

This is not to say that because all responses cannot be anticipated, they should be ignored when conducting cyber-attacks. Rather, it is to argue that the possible effects of all types of actions should be considered. There is no reason to single out cyber-attacks as being more likely than physical attacks to lead to severe retaliatory strikes with conflict escalation. Indeed, because cyber-attacks are often difficult to attribute and hence deniable by their perpetrators, they might be less likely to provoke a retaliatory strike. The targeted state may not be sure who to retaliate against, and so proceed on the side of caution rather than risking an unprovoked counter strike against innocent parties. Even if the target correctly identifies the origin of the cyber-attack, it might retaliate with an in-kind cyber-attack, which may be less harmful than had it chosen a kinetic strike. The

general non-lethal and less destructive nature of cyber-attacks gives further reason to predict that nations will respond with less damaging counter attacks, if they respond at all. Indeed, as has been noted, the very nature of cyber-attacks often makes it possible for a given target to recover from an attack quickly and with little or no permanent damage. This would allow a target nation to "save face" and either deny the attack or downplay the damage it caused.[13] Such scenarios, it seems to us, would be less likely to result in conflict escalation than comparable kinetic strikes would (Lucas 2011b; Owens et al. 2009). If not, they would at least allow for an "escape valve" to avoid direct kinetic hostilities in ways that a comparable kinetic strike would not.

Objection 6: Cyber weapons make warfare too easy

Some argue that cyber weapons makes warfare too easy. Whereas states may be reluctant to conduct physical strikes, they may be less reluctant to conduct remote attacks in cyberspace. The result could be the launching of more wars, including wars that use both cyber and physical weapons, than would otherwise happen. The fear here is not merely more wars but, most likely, when more wars are launched, there will be more unjust wars. This is the familiar "threshold" problem for all new advances in military technology (Strawser 2010; Lucas 2011a). The worry is that by making war too easy, cyber weapons will entice states to undertake war in violation of the restrictions normally imposed by the *jus ad bellum* principle of last resort.

Our response is that the principle of CyberOP protects against this and blocks the objection. It assumes that cyber weapons are used in a context where a comparable physical action is already deemed just under the traditional just war convention and LOAC. Thus, the cyber operation should not be construed as in illegal act of war violating the principle of *jus ad bellum* or *jus en bello*. It should not be seen as an illegitimate act of force or aggression, and should not lead to a full-scale, unrestricted war. Further, as was argued in Strawser (2010: 358-360), this "threshold" problem is a difficulty for every new advance in military technology; it is not unique to cyber weapons just as it is not unique to UAVs. And if a given technology is used justly in a present case and is the morally obligated weapon choice due to considerations of unnecessary risk or harm, then mere speculation about its future misuse should not trump the present normative obligation to so use it.

Notice that there is, in fact, potential for moral gain here vis-à-vis the inverse of this objection. The moral worry raised by this objection is that the ease with which a nation-state can use cyber weapons lowers the threshold to resort to war and could thereby result in more wars, which presumably means more unjust wars. But the inverse could also be true: that just causes that *should* be persecuted but are not, could be carried out by nation-states willing to use cyber-weapons who would not be willing to use kinetic weapons (even if they should).[14] In our view, this cannot stand alone as a positive

[13] This is exactly what played out with the Stuxnet attack on Iran.

[14] Savulescu and Beauchamp (forthcoming) argue for a similar moral gain that could be possibly had with regard to the increasing use of UAVs.

argument for the development and employment of cyber-weapons, because it is equally possible that they could be used for nefarious ends. But if cyber-weapons are used in line with the strictures of CyberOP, such use could result in precisely this kind of normative gain because states may be more willing to use cyber weapons due to the advantages they provide with regard to force protection through the avoidance of unnecessary risk.

In addition, we expect states following the obligations of CyberOP to be cautious about using cyber weapons, because of the uncertainty about how a target might respond. In that regard, cyber weapons serve as a deterrent against state use, not because their use would be so devastating (as with nuclear weapons), but because of the uncertainty that the target might respond in a manner that is devastating (say by using nuclear weapons). Thus, we would expect militaries to use cyber weapons more for surgical strikes conducted in the context of just wars, and for just covert operations having limited effects such as, for example, disabling local communications, power, or alarms long enough for a hostage rescue mission to complete successfully.

Conclusions

We do not claim that all cyber weapons are ethical in principle. Indeed, many are not by their very design. Rather, our claim is that in limited circumstances and granting certain assumptions, cyber weapons are morally preferred over their kinetic counterparts resulting in an obligation to use them in line with traditional just war theory principles. In particular, they are a better option when they can be deployed for a purpose already deemed just under LOAC and without any significant loss of capability. This moral preference arises out of the simple moral obligations imposed by the PURH. That is, cyber-attacks can be less risky and harmful than kinetic strikes, and can thereby impose a duty for militaries to use cyber weapons in place of their kinetic counterparts.

We leave open the ethical questions surrounding the use of cyber weapons that do not have apparent kinetic counterparts and hence are not covered by CyberOP. An example would be a cyber weapon that alters data on an adversary system so as to present false information to the adversary. Stuxnet did this. In addition to altering the code driving the centrifuges so as to physically damage them, it altered the data displayed to the operators so as to hide the effects of the attack. The ethics of this and other operations not covered by CyberOP requires additional moral reflection, including consideration of basic LOAC principles, in order to determine whether the operations are just.

We also leave open the ethics of using cyber weapons that have some of the capabilities found in kinetic weapons, but not their full capabilities. It may be that such cyber weapons are still ethically superior because of other capabilities that are not present in the kinetic weapons. The moral permissibility of such weapons would depend crucially on their ability or lack thereof to be used in a just manner as part of an overall just attack.

Although we have focused on the conduct of cyber-attacks rather than cyber exploitation (espionage), the same general reasoning might apply to exploitations. In

particular, when foreign intelligence can be collected through a cyber-operation as opposed to one that requires physical presence in foreign territory or the turning of a foreign insider (e.g., to leak classified documents), then the cyber operation might be preferred and perhaps even morally obligatory on the grounds that it would be less risky to collectors and less harmful to those collected against. However, we leave a thorough analysis of this for future study.

References

Arquilla, J. 1999. Ethics and Information Warfare. In *The Changing Role of Information in Warfare*, eds. A. Khalilzad, J. White, A. Marshall, 379-401. Santa Monica: RAND Corporation.

Broad, W. J., Markoff, J. and Sanger, D. E. 2011. Israeli Test on Worm Called Crucial in Iran Nuclear Delay. *The New York Times*, January 15.

Clarke, R. A. and Knake, R. K. 2010. *Cyber War*, New York: Harper Collins.

Denning, D. E. 2008. The Ethics of Cyber Conflict. In *The Handbook of Information and Computer Ethics*, eds. K. E. Himma and H. T. Tavani, 407-428. Hoboken: Wiley.

Dipert, R. 2012. Ethical Aspects of Cyberwar. Paper presented at the meeting on Ethical and Societal Issues in National Security Applications of Emerging Technologies for the National Academy of Sciences Division on Engineering and Physical Sciences, Computer Science and Telecommunications Board, Washington D.C.

Dipert,R. 2010. The Ethics of Cyberwar. *Journal of Military Ethics* 9(4): 384-410.

DoD. 1999. An Assessment of International Legal Issues in Information Operations. Second edition; November, Arlington: Department of Defense, Office of General Counsel.

Falliere, N., Murchu, L.O. and Chien, E. 2011. W.32 Stuxnet Dossier, V. 1.4, Symantec Security Response. February.

Frowe, H. 2011. Self-Defence and the Principle of Non-Combatant Immunity. *Journal of Moral Philosophy* 8: 530-546.

Lucas, G. R. 2011a. Industrial Challenges of Military Robots, *Journal of Military Ethics* 10(4): 274-295.

Lucas, G. R. 2011b. Permissible Preventive Cyberwar: Restricting Cyber Conflict to Justified Military Targets. Oxford: Oxford Institute for Ethics, Law, and Armed Conflict.

McMahan, J. 2009. *Killing in War*. Oxford: Oxford University Press.

Orend, B. 2000. Jus Post Bellum, *The Journal of Social Philosophy* 31(1): 117-137.

Owens, W. A., Dam, K. W. and Lin, H. S. (eds). 2009. *Technology, Policy, Law, and Ethics Regarding US Acquisition and Use of Cyberattack Capabilities*. Washington, DC: The National Academies Press.

Rattray, G. and Healey, J. 2010. Categorizing and Understanding Offensive Cyber Capabilities and Their Use. In *Proceedings of a Workshop on Deterring Cyberattacks*, eds. K. W. Dam and W. A. Owens, 77-97. Washington, DC: The National Academies Press, 77-97.

Rid, T. 2011. Cyber War Will Not Take Place. *Journal of Strategic Studies*, iFirst Article: 1-28.

Rodin, D. 2005. *War and Self-Defence*. Oxford: Oxford University Press.

Rowe, N. C. 2010a. The Ethics of Cyberweapons in Warfare. *Internationl Journal of Technoethics*, 1(1): 20-31.

Rowe, N. 2010b. Towards Reversible Cyberattacks. In *Proceedings of the 9th European Conference on Information Warfare and Security*, ed. J. Demergis, 261-267. Reading: Academic Publishing Ltd.

Savulescu, J. and Z. Beauchamp. 2013. Robot Angels: The Use of UAVs in Humanitarian Military Intervention. In *Killing By Remote Control: The Ethics of an Unmanned Military*, ed. B. J. Strawser, 106-125. New York: Oxford University Press.

Schmitt, M. N. 1999. Computer Network Attack and the Use of Force in International Law: Thoughts on a Normative Framework. *Columbia Journal of Transnational Law* 7: 885-937.

Schmitt, M. N. 2010. Cyber Operations in International Law: The Use of Force, Collective Security, Self-Defense, and Armed Conflicts. In *Proceedings of a Workshop on Deterring Cyberattacks*, eds. K. W. Dam and W. A. Owens, 151-178. Washington, DC: The National Academies Press.

Steinhoff, U. 2013. Extreme Asymmetry and its Discontents. In *Killing By Remote Control: The Ethics of an Unmanned Military*, ed. B. J. Strawser, 179-207. New York: Oxford University Press.

Strawser, B. J. 2010. Moral Predators: The Duty to Employ Uninhabited Aerial Vehicles. *Journal of Military Ethics* 9(4): 342-368.

Symantec. 2011. Internet Security Threat Report, Trends for 2010, Vol. 16, April.

Tikk, E., Kaska, K., and Vihul, L. 2010. *International Cyber Incidents: Legal Considerations*. Estonia: NATO Cooperative Cyber Defence Centre of Excellence.

Wingfield, T. 2000. *The Law of Information Conflict*. Falls Church, VA: Aegis Research Corporation.

Wingfield, T. 2009. International Law and Information Operations. In *Cyberpower and National Security*, eds. F. D. Kramer, S. H. Starr, and L. K. Wentz, 525-542. Washington, DC: NDU Press.