Cyberwar and Mediation Theory

*Nolen Gertz, Peter-Paul Verbeek and David M. Douglas**

Cyberwar (military operations conducted via computer networks) is often downplayed compared to traditional military operations as they are largely invisible to outside observers, difficult to convincingly attribute to a particular source and rarely cause physical damage or obvious harm. We use mediation theory to argue that cyberwar operations cause harm by undermining trust in computerised devices and networks and by disrupting the transparency of our usage of information technology in our daily lives. Cyberwar operations militarise and weaponise the civilian space of the Internet by co-opting and targeting civilian infrastructure and property. These operations (and the possibility of such operations occurring) fundamentally change users' Internet experience by fostering fear and paranoia about otherwise unnoticed and transparent aspects of their lives, similarly to how biological and chemical weapons create fear and paranoia about breathing, eating, and physical exposure to the world. We argue that the phenomenological aspects of cyberwar operations offer a compelling justification for prohibiting cyberwar in the same manner in which biological and chemical warfare are prohibited.

I. What the Cyberwar Debate Reveals about our Views of Cyberspace

In response to the confusion surrounding the militarisation of cyberspace, Taddeo[1] has tried to bring together the traditional just war theory (JWT) and the revolutionary information ethics (IE) to create a new framework for understanding cyberwarfare. This new framework, which Taddeo has named 'just information warfare' (JIW), is an attempt to merge the principles for adjudicating warfare found in JWT with the 'ontocentric' ethics of IE. Whereas the principles of JWT such as last resort, proportionality, right intention, legitimate authority, and discrimination are meant to be used to determine if a war is or is not justified,[2] JWT is a traditional and anthropocentric theory. For this reason trying to simply apply JWT to cyberwarfare leads, as Taddeo argues, to confusion over how to apply traditional thinking about war to a non-traditional battlefield like cyberspace.

IE is neither traditional nor anthropocentric, as it is based on an ontology of information, an ontology that sees reality as an 'infosphere' populated, not by human beings and non-human beings, but by 'informational beings'.[3] By moving from a dualistic to a monistic ontology, Taddeo is able to fill in the gaps that appear when trying to apply traditional theories of warfare to cyberwarfare, particularly the gap of how to ethically account for the damage and destruction of cyber-based targets. This is achieved by focusing not on protecting human life as a *casus belli*, but rather on protecting the infosphere:

DOI: 10.21552/delphi/2019/2/5

* Nolen Gertz, Assistant Professor of Applied Philosophy, University of Twente. For correspondence: <n.gertz@utwente.nl>. Peter-Paul Verbeek, Professor of Philosophy of Technology, University of Twente. For correspondence: <p.p.c.c.verbeek@utwente.nl>. David M Douglas. For correspondence: <dmdouglas256@tuta.io>.

1 Mariarosaria Taddeo 'Just Information Warfare' (2014) Topoi 35, 213-224

2 These and other principles that can be found in JWT have a lengthy history of debate behind them that stretches from the Peloponnesian War to today. The principles were arrived at by philosophers and theologians who were concerned with trying to help political and military leaders to determine when and how it is 'just' to fight a war. JWT can therefore be seen as an attempt to avoid what theorists in this tradition see as the dangers of the murderous history of political realism and of the suicidal defeatism of pacifism.

3 Luciano Floridi, *The Ethics of Information* (OUP 2013). According to Floridi, humans and websites, animals and computer programs can all be seen as information – think for example of the parallels between a human's DNA sequence and a website's HTML code. Floridi argues on the basis of this 'infocentrism' that all such informational beings or 'clusters' deserve to be respected, to be accorded some amount of moral value, based on an informational hierarchy.

> ...an entity may lose its rights to exist and flourish when it comes into conflict (causes entropy) with the rights of other entities or with the well-being of the Infosphere. It is a moral duty of the other inhabitants of the Infosphere to *remove* such a malicious entity from the environment or at least to impede it from perpetrating more evil.[4]

JIW appears to clear up the blind spots and confusion of traditional theories, but at the cost of losing one of the major principles of JWT: discrimination. The principle of discrimination requires that militaries observe a strict distinction in targeting between combatants and noncombatants. However, in cyberspace it is much more difficult to make such distinctions, not only because any user is a potential 'malicious entity' – IW rightfully targets only malicious entities, be they military or civilian',[5] but because of the very nature of cyberspace.

This inability to maintain the principle of discrimination in cyberwarfare is typically referred to as the 'attribution problem.'[6] In traditional warfare, discrimination is maintained by requiring that militaries wear uniforms and fight as far from civilian populations as possible, both of which seem to be impossible requirements to maintain in cyberspace. While there have been attempts to address this issue, such as using a broader distinction between 'licit targets and illicit ones'[7] or calling for an international agreement to only attack using 'digital signatures'[8] these approaches do not touch on the larger issue of what it means to turn a predominantly civilian space into a battlefield, to weaponise and militarise cyberspace. Taddeo[9] believes that cyberwarfare can be justified so long as it is only intended to return the infosphere to its *status quo ante*, but this requires that we know what the *status quo ante* is.

Hence without a better understanding of the implications and consequences of cyberwarfare for both military and non-military users, we cannot successfully justify or deter cyberwarfare. It is for this reason that we argue that rather than trying to understand cyberwarfare through the lens of IE, we should instead turn to mediation theory, which can help us to analyse the human-technology relations involved in cyberspace and cyberwar. And this will make it possible to show that the cyber/physical and real/virtual dualisms underlying JIW fail to recognise that cyberwarfare technologies are not merely operating in the 'virtual' realm of cyberspace, but play an important normatively-mediating role in the 'real world' as well.

II. Mediating Cyberspace

Mediation theory is a descriptive and normative theory of human-technology-world relations, of how technologies mediate the relationships humans have to the world. While its descriptive side is a continuation of postphenomenology, its normative side is an attempt to use postphenomenology to reveal how technologies *already* mediate ethical life, and to decide how technologies *should* mediate ethical life. If, as Ihde[10] shows, technologies mediate human experience in the form of embodiment, hermeneutic, alterity, and background relations, then, as Verbeek[11] argues, humans are not autonomous in the way that traditional ethical theories suppose, as technologies do not only participate in, but actively shape ethical decisions and actions.

Analysing cyber technologies through the lens of mediation theory makes it possible to show that these technologies in fact cannot be understood as being part of a 'virtual reality'. Rather, they need to be understood as mediators in the real world. First, the virtual world that is opened up by cyber technologies is a world that is experienced by 'real humans', via keyboards, screens, VR glasses, cameras, speakers, and microphones. And therefore, second, what happens in the virtual realm will inevitably mediate what happens in the real world: ultimately, it does not happen to virtual beings, but to real beings, whose avatars and virtual words are part of their lifeworld.

4 (n 1)

5 (n 1)

6 Randall R Dipert, 'The future impact of a long period of limited cyberwarfare on the ethics of warfare' in Luciano Floridi and Mariarosaria Taddeo (eds), *The Ethics of Information Warfare* (Springer 2013) 25-38; Kenneth Geers 'The Challenge of Cyber Attack Deterrence' (2010) Computer Law & Security Review 26, 298-303

7 (n 1)

8 Patrick Lin et al, 'Is warfare the right frame for the cyber debate? Luciano Floridi and Mariarosaria Taddeo (eds), *The Ethics of Information Warfare* (Springer 2013) 39-60

9 (n 1)

10 Don Ihde, *Technology and the Lifeworld* (Indiana University Press 1990)

11 Peter-Paul Verbeek, *Moralizing Technology* (The University of Chicago Press 2011)

To carry out a mediational analysis of cyberspace, we must start by recognising that cyberspace is not one technology, but rather a complicated complex of many interconnected technologies. We might say for example that we are using our laptops to get online, but in reality we are using at least a keyboard, a mouse, and a display to use programs that are graphical user interfaces interpreting and translating codes and scripts sent by us and to us over an internet connection. That we do not refer to each of these processes, processes which of course could be even further broken down and described in ever greater detail, is indicative of the relationship between users, computers, and cyberspace. Rosenberger,[12] following Ihde, describes this as 'the deeply embodied character of human relations to technologies such as the computer.'

Embodiment relations, according to Ihde,[13] occur when technologies become an extension of the user, enabling the user to carry out a project through the technology, a project that could not have been carried out without the transformative mediation of the technology. A key aspect of this relation is that we focus on *what we achieve through the technology*, rather than on *the technological mediation itself*, such as when I say I see something rather than saying I see something through my glasses, or I say I go online rather than saying I go online through my computer. This technological 'transparency'[14] is not a side-effect of using a technology such as a computer, but is rather essential to its use, for the more aware we become of the mediating role of the technology, the less the technology can affect the mediation, requiring that we focus not on our ends but instead on the means to achieve them.

As Heidegger first pointed out, when technologies cease to function properly they lose not only their functionality, but their transparency, turning from the invisibility of being 'ready-to-hand' to the obtrusiveness of being 'present-at-hand.' The malfunctioning technological artefact is no longer an extension of one's embodiment, existing instead as a mere *thing*, calling attention to not only its own limitations, but to the limitations of the user as well. Though Heidegger uses a broken hammer as an example, this relational breakdown can also apply to a computer:

> The user's everyday relationship with the computer is disrupted when it acts in an unexpected way. Suddenly, aspects of computer use that had faded into the background explode again into the foreground. The experience can be jarring. The user becomes plainly aware that the options for interacting with the technology are limited, that particular keys can be pressed, that the computer responds to some things and not others. Though it had been an invisible tool a moment ago, the computer now sits as an obstacle between one and one's work. The user again becomes aware of her or his place in front of the device. When a computer unexpectedly and abruptly ceases to work properly, a user may become explicitly conscious of the computer's identity as a technology, and of her or his situation as a user, all of the sudden.[15]

When the computer functions as expected, we lose sight of it, working not *at* it or *on* it, but *with* it and *through* it. It is for this reason that we tend to use spatial metaphors when referring to the internet, as we experience not code on a computer screen, but a cyber*space*, where we can *go* online, *surf* the web, and *follow* other users on social media, often for much longer than we realise, as we lose sight of ourselves as well. At the same time, the screens, keyboards, and other devices we embody help to shape *how* we go online, *what* cyberspace means, and *how* 'following' becomes a dimension of sociality.

When the computer functions in an unexpected way however, we regain sight of it and of ourselves, experiencing the computer no longer as an 'invisible tool,' but as 'an obstacle between one and one's work.' To describe the malfunctioning computer as an 'obstacle' is to recognise that, while it may appear that there is something to be gained from a malfunction forcing one to pay more attention to one's computer and to one's use of it, this increased attention is often experienced not as a discovery, but as a distraction. A malfunction does not tend to lead us to new insights and learning opportunities, as instead we become irritated, angry, fixated, unable to focus on anything or anyone other than the malfunction, for which reason, as Rosenberger[16] puts it, 'the experience can be jarring.' Part of what is so 'jarring' about

12 Robert Rosenberger, 'The Sudden Experience of the Computer' (2009) *AI & Society* 24, 173-180

13 (n 10)

14 (n 10); (n 12)

15 (n 12)

16 (n 12)

such an experience is that the malfunction moves the computer along 'the human-technology continuum of relations' from the one end of embodiment relations to the opposite end of 'alterity relations'.[17] Ihde writes:

> Word processors have become familiar technologies, often strongly liked by their users... Yet in breakdown, this quasi-love relationship reveals its quasi-hate underside as well. Whatever form of 'crash' may occur, particularly if some fairly large section of text is involved, it occasions frustration and even rage. Then, too, the programs have their idiosyncrasies, which allow or do not allow certain movements; and another form of human-technology competition may emerge. (Mastery in the highest sense most likely comes from learning to program and thus overwhelm the machine's previous brainpower. 'Hacking' becomes the game-like competition in which an entire system is the alterity correlate.) Alterity relations may be noted to emerge in a wide range of computer technologies that, while failing quite strongly to mimic bodily incarnations, nevertheless display a quasi-otherness within the limits of linguistics and, more particularly, of logical behaviors. Ultimately, of course, whatever contest emerges, its source lies opaquely with other humans as well but also with the transformed technofact, which itself now plays a more obvious role within the overall relational net.[18]

As Ihde makes clear, a malfunctioning computer can present itself as more than just an obstacle, as it can become a competitor, an object of 'quasi-hate' that, as Ihde goes on to point out, can provoke fantasies opposite to those of embodiment relations. With alterity relations we no longer dream of becoming one with the technology, but instead have, as we see again and again in pop culture, the fears that the 'brain power' of computers would soon replace human thinking, fears that political or military decisions will not only be informed by but also made by computers'.

However, a malfunction does not only impact our embodiment relations to our computers as a key element of our relationship to computers, particularly with regards to cyberspace, can be described as a 'hermeneutic relation'.[19] When we say that we explore cyberspace, go online, surf the web, and follow social media, what such shorthand is most often referring to is the act of reading. Similar to embodiment relations, the hermeneutic relation of reading requires transparency, not only of the technological device conveying the texts and images we read, but of the texts and images themselves. In reading we lose not only ourselves, but also the lines that make up the letters that make up the words that make up the sentences of the book or web-pages before our eyes. This 'hermeneutic transparency'[20] allows us to experience what we are reading in an embodied manner, enabling us to project ourselves into what we are reading without the distraction of being aware that we are interpreting symbols, moving our eyes, and turning pages with our hands.

What we are also unaware of due to hermeneutic transparency is the faith and trust we put into what we are reading. Hermeneutic relations are, like embodiment relations, a form of technological mediation, but unlike embodiment relations, hermeneutic mediations present the world to us rather than helping us to extend our bodies towards the world. The danger of hermeneutic relations therefore is not only that of misplaced trust, for, as Ihde[21] points out with regards to the example of the Three Mile Island disaster, being misled by what one is reading can have life and death consequences. Merely being made aware of the possibility of hermeneutic breakdown, of the possibility of reading something that does not reflect the truth of the world it purports to show us, can be enough to make us paranoid and make the transparency required for reading nearly impossible. For technologies to not function as expected with regards to hermeneutic relations leads, as it does with regards to embodiment relations, to being forced to become aware of the mediating technology and our dependence on it, but it is much easier to replace a malfunctioning computer than it is a malfunctioning news source. Paranoia that one's new computer will eventually break just like the previous computer may lead to irritation, but paranoia that one's new source of information will eventually mislead just like the previous source may lead to a general mistrust of the medium itself, as for example has hap-

17 (n 10)

18 (n 12)

19 (n 10)

20 ibid

21 ibid

pened with both Wikipedia and the 'mainstream media' having become euphemisms in the United States for untrustworthy information.

The fourth type of human-technology relations that has a central place in Ihde's typology is the background relation.[22] Here, technologies are not experienced or embodied directly, but form the context of our experiences. The air conditioning system that is making noise all the time. The notifications our mobile devices are giving us all day, to inform us about messages and calls we receive, and the activity of other people on social media. In fact, many information technology systems have started to form the background of our daily lives. And by contextualising our existence, these technologies have a profound influence on the ways in which we live our lives – up to the point that they are in fact conditioning it, from their roles at the background. Banking systems, communication systems, traffic control systems, surveillance systems are not just neutral backgrounds, but shape the character of our daily lives by forming its context.

Beside analysing the specific relations that can arise between humans and technologies, mediation theory also addresses the *normative* dimension of these relations. When technologies help to shape our interpretations of the world, and the actions and practices we engage in, they in fact help us to do ethics: they inform our moral decisions and actions. MRI imaging mediates moral decisions about the lives of coma patients,[23] just like drones mediate the moral engagement of operators with their victims.[24] Technologies mediate morality - they have become part of the moral agency of human beings. Moreover, technologies help to shape normative frameworks. Our norms regarding acceptable forms of suffering, for instance, have developed in interaction with anesthetic technologies. While anesthesia used to be highly contested in its early days, it has now become immoral to operate on people without giving them proper anesthesia. Technologies are morally neutral, but help us to do ethics by informing moral actions, decisions, and frameworks. From the perspective of mediation theory, then, cyber technologies should not be analysed and evaluated as technologies that primarily affect the virtual realm. As we saw, information technologies play a profoundly mediating role in our 'real' lives as well. By being embodied, read, and interacted with, and from their role as the contextual background of our lives, cyber technologies have become an integral part of what it means to be human in a digital era. Moreover, these technologies help to shape the moral decisions we make and the moral frameworks from which we think.

III. Mediating Cyberwar

Now that mediation theory has helped to clarify the conceptual confusion surrounding cyberspace, we can turn our focus to the conceptual confusion surrounding cyberwar, and in particular on what it means to weaponize and militarise cyberspace. If, as we have seen, cyberspace is not a virtual realm independent of the real world, but rather is composed of technologies that are essential to the fabric of our daily lives, then cyberwarfare should be thought of as being on par less with crime and espionage[25] and more with chemical and biological warfare.[26] Deterrence of cyberwarfare must therefore be sought not through regulation[27] but through international treaties that ban the practice altogether.

Mediation theory allows us to highlight an issue for deterrence against cyber attacks, which is how the capability to employ active defenses or to retaliate transforms the Internet from a medium that is benign in itself to a militarized medium that may be employed to cause harm. This becomes apparent when we consider that many of the tools of cyber attacks are dual-use technologies that are well within the means of the technologically-savvy to acquire and use. This is another aspect of the attribution problem that makes it difficult to distinguish the actions of states and non-state actors. If a non-state actor could launch a cyber attack, another non-state actor may retaliate in kind. The Internet would become an avenue through which individuals, either unwillingly or unknowingly, may become targets of states who have identified them as attackers.

---

22 ibid

23 (n 11)

24 Nolen Gertz, *The Philosophy of War and Exile* (Palgrave-Macmillan 2014)

25 Patrick Lin et al, 'Is warfare the right frame for the cyber debate? Luciano Floridi and Mariarosaria Taddeo (eds), *The Ethics of Information Warfare* (Springer 2013) 39-60

26 Gregory Koblentz and Brian Mazanec, 'Viral Warfare: The Security Implications of Cyber and Biological Weapons' (2013) Comparative Strategy 32(5), 418-434

27 (n 1)

The possibility of being targeted by states may lead individuals to reconsider how they use the Internet, and as a result, change the Internet itself. Jonathan Zittrain[28] writes that an important characteristic of the Internet is its *generativity*: the 'capacity to produce unanticipated change through unfiltered contributions from broad and varied audiences.' This is in contrast to what he[29] calls 'sterile' systems that are under strict control that limits their ability to be used for unanticipated uses. This difference is best illustrated by comparing the current Internet to the network services of the 1980s and 1990s, such as CompuServe and AOL (America Online). Only software supplied by the service operators could access these services, and the uses of these services were strictly controlled by their operators. While such services could not be used as sources for cyber attacks, this level of security would seemingly result in a Catch-22 type paradox: the Internet could be a protected environment, but at the cost of destroying everything that makes cyberspace worthy of being a protected environment.

IV. Conclusion: Regulating vs Banning Cyberwarfare

In 1969, President Richard Nixon unilaterally renounced the possession and use of biological weapons by the United States. As he explained, 'These important decisions have been taken as an initiative towards peace. Mankind already carries in its own hands too many seeds of its own destruction. By the examples we set today, we hope to contribute to an atmosphere of peace and understanding between nations and among men.'[30] Shortly thereafter, in 1972, the Biological Weapons Convention (BWC) was opened for signature in London, Moscow, and Washington, with 110 states becoming signatories and 173 states becoming parties to the BWC. The BWC categorizes biological weapons as 'weapons of mass destruction' and states that the use of biological weapons 'would be repugnant to the conscience of mankind and that no effort should be spared to minimise the risk.'[31]

The moral arguments that motivate this rejection of biological warfare are grounded in 'its uncontrollable and indiscriminate effects, its insidious nature, and its deliberate perversion of medical science.'[32] Koblentz and Mazanec outline[33] the following parallels between biological weapons and cyber weapons:

> Both types of weapons pose significant challenges to attribution; are attractive to weaker powers and nonstate actors as an asymmetric weapon; have the potential for use as a force multiplier for conventional military operations; are of questionable deterrent value; exhibit a high degree of uncertainty and unpredictability in their use and the potential for major collateral damage or unintended consequences; are based on multi-use nature of the underlying technologies; and are typically developed under highly secretive programs.

On the basis of these parallels, they conclude that cyberwarfare should not be regulated, but banned, with 'dissuasion as the heart of a long-term strategy for managing the risks posed by cyber and biological weapons.' Similar to the BWC, successful cyberwarfare dissuasion would require 'a widely shared understanding among nations and societies that advances in information technology should only be used for peaceful purposes and the use of cyber weapons to attack civilian targets and critical infrastructure is unacceptable.'[34] However, as Koblentz and Mazanec point out, it is easier to dissuade states from acquiring and using biological weapons as there is already a 'taboo associated with poisons,' while 'cyber weapons, as relatively novel creations that operate in a new and man-made domain, lack such a similar historical, normative framework.'[35]

Mediation theory has helped us to see that while there are indeed discrepancies in our thinking about biological and cyber weapons, one reason for such discrepancies is the perpetuation of dualistic thinking about cyberspace, in particular that cyber weapons 'operate in a new and man-made do-

28 Jonathan Zittrain, *The Future of the Internet* (Penguin Books 2008)

29 ibid

30 Jonathan B Tucker Tucker, 'A Farewell to Germs' (2002) International Security 27(1), 107-148

31 United Nations Office for Disarmament Affairs, 'Convention on the prohibition of the development, production and stockpiling of bacteriological (biological) and toxin weapons and on their destruction' (2016) <http://disarmament.un.org/treaties/t/bwc/text> accessed 31 July 2016

32 Brian Balmer, 'Killing 'Without the Distressing Preliminaries': Scientists' Defence of the British Biological Warfare Programme' (2002) Minerva 40(1), 57–75

33 (n 26)

34 ibid

35 ibid

main'.[36]Cyber technologies – through forming embodiment, hermeneutic, and background relations with users – belong to the same domain as biotechnologies, the domain of everyday life. It is for this reason that the weaponisation and militarisation of cyberspace is the weaponisation and militarisation of everyday life.

A further parallel between biological and cyber weapons is thus that of *paranoia*. Releasing toxins into the atmosphere can not only result in 'uncertain area coverage and effects' due to 'environmental and meteorological conditions'[37] but fear and anxiety about being able to safely go outside. Mediation theory reveals how cyberspace can better be thought of as part of the atmosphere we breathe rather than as a tool we use only on occasion, for which reason the proliferation of cyber weapons can create a paranoia of the space around us, of the space we live in, just as much as can biological weapons. Gas masks and hazmat suits allow us to enter regions infected by biological weapons, but do so at the cost of requiring that the wearer be limited in activities and disrupted in one's tasks through the diminished freedom and constant awareness of the protective gear. Antivirus software, firewalls, and anti-surveillance hardware similarly allow us to enter regions infected by cyber weapons, but again at the cost of requiring the user be limited and disrupted by the diminished freedom and constant awareness of cyberprotection.

If Koblentz and Mazanec[38] are calling for a ban of cyber weapons on the basis of parallels with biological weapons, while yet maintaining dualistic thinking about cyber weapons, then mediation theory, by removing such dualistic confusion, only further strengthens the argument in favor of a cyber weapons ban. Such a ban would of course not make it impossible for states to still operate cyber weapons research programs, much like how the BWC did not prevent the Soviet Union from maintaining a secret biological weapons program. However, denouncing cyber weapons and stigmatising their use would help to create new norms and strengthen existing norms of safe Internet use, as well as help to expand our understanding of the interconnected nature of cyberspace. Attempts to regulate cyberwar will create the false impression that there are safe and manageable ways to weaponise and militarise cyberspace. A cyber weapons ban however would lead more and more people to realise that what happens in cyberspace doesn't stay in cyberspace, which would in turn persuade more and more people to protect rather than risk endangering the cyber backbone of our daily lives.

36 ibid
37 ibid
38 ibid