The evolution of targeted killing practices: Autonomous weapons, future conflict, and the international order

Michael Carl Haas & Sophie-Charlotte Fischer

To cite this article: Michael Carl Haas & Sophie-Charlotte Fischer (2017) The evolution of targeted killing practices: Autonomous weapons, future conflict, and the international order, Contemporary Security Policy, 38:2, 281-306

To link to this article: http://dx.doi.org/10.1080/13523260.2017.1336407
The evolution of targeted killing practices: Autonomous weapons, future conflict, and the international order

Michael Carl Haas and Sophie-Charlotte Fischer

Center for Security Studies, ETH Zurich, Zurich, Switzerland

ABSTRACT
This article examines the potential use of autonomous aerial weapons for targeted killing purposes and, in doing so, looks beyond the now-familiar “global war on terror.” We argue that the combination of novel capabilities with the pre-existing military-theoretical frameworks of advanced Western states, within which autonomous weapons will be embedded, may be conducive to an expansion of targeted killings to scenarios other than military counter-terrorism. The confluence of autonomous weapons and targeted killing practices may therefore lead to a further weakening of long-standing norms regulating the use of force, including in interstate scenarios. We also find that international regulation is unlikely to forestall this outcome, and that political-military insistence on centralized operational control may mitigate—but not negate—the disruptive potential of these developments. As a result, the possible consequences for the international order of an evolution of targeted killing practices along these lines should not be underestimated.

KEYWORDS Aerial warfare; autonomous weapon; international order; interstate war; hybrid war; targeted killing

The early 21st century has seen an increasing reliance on targeted killings as a means of international security policy. This “normalization of individualized targeted killing as a practice used by the state” (Waldron, 2015, p. 292) was, by and large, a state reaction to the emergence of pervasive terrorist threats, which were seen as emanating from territorial “safe havens” and directed by centralized leadership structures (Johnston, 2012; Jordan, 2014b; Krishnan, 2013; Price, 2012). Significant policy and concomitant legal shifts have enabled this massive expansion of the use of targeted killings, as have perceived operational requirements of a more ad hoc nature. Many of these developments are addressed at length in this special issue. But policy and exigency have not been the only “game changers” in this regard. As Senn and Troy (2017, p. 190) emphasize in their introduction, technological change also matters. To this day, targeted killings can be of an intensely up close
and personal nature, as is attested to by the killing of Osama Bin Laden by U.S. Navy special operations forces in 2011 and hundreds of other recent kill/capture operations (Naylor, 2015, pp. 391–402). It is equally apparent, however, that the technological revolutions of the 20th century have had a profound impact on the application of this controversial method, which has long been spurned by the great majority of states.

This article investigates the potential of 21st-century advances in technology to revolutionize targeted killings to a similar, or perhaps an even greater, extent than the late 20th-century revolution in precision-guided munitions (PGMs) has done (Watts, 2007). Specifically, we ask what impact the possible introduction of autonomous weapons is likely to have on the scope and frequency of targeted killings in future conflicts. Our focus in addressing this question is on one main strand in the evolution of targeted killings, which we believe will continue to play a particularly prominent—perhaps even dominant—role: the reliance on aerial platforms and weaponry to conduct such killings from afar.

We develop our argument in four steps. In the first part, we provide an overview of emerging developments in machine autonomy, as they relate to a potential future expansion of targeted killing practices. In the second part, we examine the military-theoretical frameworks within which autonomous weapons will necessarily be embedded. We then begin to move the targeted killing debate beyond the now-familiar context of the “global war on terror” and examine the potential use of autonomous aerial weapons in future conflict scenarios, which we believe may extend well beyond its current scope of military counter-terrorism operations. In the final part of the paper, we look at potential constraints upon such an expansion, including possible regulatory measures, as well as military-operational and political restraints that may apply. We conclude the paper with a brief review of the implications of future targeted killing practices for an international order that is already buckling under the pressure, as Senn and Troy (2017, pp. 189–191) suggested in their introduction.

Our main findings are twofold. First, and most importantly, there is a reason to expect that targeted killing practices will be de-linked from current military counter-terrorism operations and that machine autonomy is going to contribute to this trend. Both future conventional and hybrid conflict scenarios may see the use of targeted killings for purposes including enhanced conventional deterrence, micro-targeting of enemy leadership personnel, intimidation of enemy leaders and populations, attempts at conflict limitation and escalation control, and deniability of semi-covert air operations. Many of these applications could contribute to a further erosion of existing norms on the use of force, and thus of the international order at large. While regulation efforts are not necessarily doomed to fail, we expect that they will have only limited impact, thus leaving it largely to individual
state actors to implement restraints on their respective targeted killing practices. Some of these restraints will almost certainly result from the reluctance of military and political leaders to surrender tight hierarchical control over a highly sensitive category of operations. It is unlikely, however, that such operational restraints will, by themselves, be sufficient to avert further damage to the incumbent international order, as far as its ability to effectively regulate the use of force is concerned.

The coming transformation of aerial warfare

From the U.S. Army Air Forces’ shoot-down of Adm. Yamamoto Isoruku in 1943 to Russia’s 1996 assassination of Chechen separatist leader Dzhokhar Dudayev, airpower has often been at the forefront of the evolution of targeted killing practices in the 20th century. The wave of targeted killings that began in the early 2000s has harnessed, and made liberal use of, a variety of aerial platforms and PGMs first fielded in the 1980s and 1990s, often for different purposes entirely (Lambeth, 2000, pp. 54–102). In their introduction, Senn and Troy (2017, p. 190) have already highlighted the role unmanned aircraft, or aerial “drones,” have played in the most recent wave of targeted killings. The evolution of drones for information, surveillance, target acquisition, and reconnaissance (ISTAR) as well as attack missions and the availability of advanced munitions have been crucial enablers of targeted killing practices in the “global war on terror” (Byman, 2013; Callam, 2010). Together, the means and methods of wide-area reconnaissance and precision attack from a distance have made it much easier to execute persistent campaigns of targeted killings across a large swath of territory, as opposed to the isolated attacks of the past.

We contend that this pattern of adapting emerging means and methods of aerial warfare to a targeted killing “mission set” is likely to persist, for as long as a number of states with advanced air forces—including the United States, the United Kingdom, and Israel—routinely conduct such operations. Therefore, we argue that important clues as to the future of the practice of targeted killing can be derived from an analysis of ongoing trends in military affairs, which are likely to be embraced by the actors at the helm of the most recent expansion of targeted killing.

Currently, the technological development with the greatest transformative potential in the field of aerial weapons is the prospect of increasing the autonomy of critical functions in weapon systems, which is made possible by accelerating advances in robotics, artificial intelligence (AI), and machine learning. Together, these advances have the potential to effect yet another revolution in aerial warfare. In the following section, we proceed to outline the observable technological trends in some detail and explain how they relate to the targeted killing “mission set.”
The rise of machine autonomy

In the course of the “war on terror,” unmanned aerial vehicles provided Western states the opportunity to respond to security threats in politically and geographically challenging environments. They enabled military commanders to maintain a close watch over potential targets from afar and allowed them to strike those targets at a moment’s notice, without a need to put soldiers’ lives at risk, or even to commit additional resources (Mayer, 2015; Sauer & Schörnig, 2012; Shaw, 2013). Data gathered by drones equipped with high-end sensors provided crucial information about the location, activities, and organizational relationships of enemy operatives. Missions of much longer duration than those typical of manned aircraft could be flown on a routine basis. Close observation of a target with no risk to friendly personnel and the use of payloads such as the air-to-ground Hellfire missile, with a warhead one-tenth the size of a 500-pound general-purpose bomb’s, could—and probably did—reduce civilian casualties to a significant extent (Byman, 2013).

As a result, relatively primitive drone technology has already changed the face of warfare, both in acknowledged conflict zones, such as Afghanistan and Iraq, and in largely unacknowledged “shadow wars” (Axe, 2013) in Pakistan, East Africa, and Yemen. Looking beyond this now-familiar context, the utility of unmanned combat aerial vehicles (UCAVs) that have seen frequent use in military counter-terrorism, such as the United States’ MQ-9 Reaper, is limited by their low airspeed, limited payloads, and lack of survivability. They also feature low degrees of automation (or none at all). While they are remotely piloted, a sizable ground crew is required to control every aspect of their operations by means of a satellite link or line-of-sight antenna (Gilli & Gilli, 2016). This includes not just drone operators and sensor specialists, but also groups of analysts, who interpret and analyze the copious amounts of data produced by the drones’ sensors. As a result, this first generation of UCAVs is now far behind the curve of burgeoning developments in military technology and will begin to be replaced by more advanced systems in the 2020s.

In fact, rapid technological advances in the fields of robotics, AI, and machine learning have already introduced a new generation of robotic systems into the civilian sphere. These same advances are being harnessed to develop robotic systems for military purposes that will feature increasing degrees of autonomy with regard to critical systemic functions (Work & Brimley, 2014, pp. 22–25). While the focus of this article is on aerial systems, the same will likely be true with regard to many military missions in the sea, land, space and cyber domains as well. These ongoing trends have the potential of cumulating in the development of autonomous weapon systems (AWS) that are distinguished by their ability to “once activated, select and engage targets without further intervention by a human
operator” (Department of Defense, 2012, p. 13). If developed and deployed, some experts believe that these systems could have a transformative impact on established methods of warfare (Work & Brimley, 2014, pp. 7–9). As far as the future use of targeted killings is concerned, we believe that the introduction of some elements of machine autonomy is highly likely even in the near term. However, to make sense of these developments, it is first necessary to gain a more nuanced understanding of the concept of autonomy than the above definition can provide.

**A functional approach to autonomy in weapon systems**

Although there has been a noticeable increase in public attention to, and academic publications on, the implications—legal and ethical, as well as strategic—of the development and deployment of AWS, the concept of “autonomy” remains poorly understood. To this day, neither academics nor policy makers have been able to agree on a common understanding of AWS, thereby complicating discussions on the potential benefits and risks associated with this important technological trend. According to Scharre and Horowitz (2015, p. 5), existing definitions emphasize one of three different dimensions of autonomy: the human–machine command-and-control relationship, the complexity of the decisions made by the machine, or the functions of the system that are made autonomous.

A frequent approach, used by actors as varied as Human Rights Watch and the U.S. Department of Defense (DoD), is to define autonomy in terms of human involvement in the execution of a machine’s tasks. Systems that are marked by full human control over all the machine’s core tasks are considered “human-in-the-loop” systems. When systems can operate independently, but humans can still review its decisions and intervene in the case of malfunction, they are often called “on-the-loop” or “human-supervised” systems. A system that can carry out its tasks completely independently, leaving no opportunity for a human to intervene, is considered an autonomous weapon or “human-out-of-the-loop” system (Scharre & Horowitz, 2015, p. 6).

However, this classification does not capture another relevant—and much more fine-grained—dimension of autonomy: namely, the complexity of machines’ decision-making processes. This dimension emphasizes the capability of a machine to incorporate uncertainties in its environment into the decision-making process and adapt to them (Boulanin, 2016a, p. 3). While systems that are classified as automatic, such as landmines, simply respond more or less mechanically to some well-defined input (Scharre, 2015, p. 10), the distinction between automated and autonomous weapons is more difficult. Often, missile defense systems such as Israel’s Iron Dome are considered highly automated as they are pre-programmed rule-based systems, which provide largely predictable outcomes (Scharre, 2015, p. 10).
Truly autonomous systems are far more complex, as they would be able to reason probabilistically based on a set of inputs, compose different courses of actions, and then select and execute the best option without human intervention at any of these stages (Cummings, 2017, p. 3).

A third, “functional view” of autonomy, which is elaborated by Boulanin (2016a) and which we will build on in this paper, emphasizes that autonomy is not a general characteristic of a system as a whole but that it matters, which decisions or tasks are completed autonomously by a machine. Therefore, autonomy needs to be understood and analyzed as an attribute that can be “attached” to different tasks of a weapon system. This concerns five core functions of weapon systems in particular: mobility, health management, interoperability, battlefield intelligence and, finally, the use of force (Boulanin, 2016a, p. 7).

Most critical with respect to the characteristics of autonomy emphasized in this paper are ISTAR and the use of force, as these functions allow the system to identify, track, select, and engage targets and to do so without direct human control at the tactical level. In other words, such a system will be fed with a set of human-approved operational parameters, within which it may seek out and attack targets as its AI sees fit, with no need for direct supervision at any point between the generation of the mission data and the system’s return to base (or, in the case on an expendable munition, its own destruction upon execution of an attack). It is systems that are autonomous in this particular sense that will have the greatest transformative impact on the practice of targeted killings and it is those systems that we will focus on in the following. The next section seeks to provide a basic understanding of the technological developments that are driving the development and future deployment of platforms and weapons that are autonomous in the performance of their ISTAR and use of force functions.

**Technologies enabling machine autonomy**

Although no truly autonomous systems—according to the above definition—have been introduced so far, the fast pace of technological development is constantly opening up opportunities for incorporating higher degrees of autonomy into the different functions of weapon systems. Notably, many technologies in the field of robotics and computer science that are key in the creation of autonomy in weapon systems are “dual-use” and driven by technological developments and applications in the commercial sector, such as the information and communications or automotive industries (Boulanin, 2016b). The foundation of machine autonomy is AI, which is primarily a “software endeavor” (Boulanin, 2016a, p. 15; Defense Science Board, 2016). Broadly defined, “AI is the capability of machines to perform complex tasks such as decision-making and perception that were normally restricted to
humans” (Cummings, 2017, p. 2). Thus, in an intelligent system, AI mimics human cognitive functions in solving specific problems and tasks (Defense Science Board, 2016). However, a further distinction is often drawn between “narrow” (or “weak”) and “general” (or “strong”) AI. While “narrow” AI refers to a system with a distinctly limited range of cognitive abilities, “general” AI denotes the capability to perform a broad range of complex intellectual tasks (Bostrom, 2014, p. 19). So far, most existing applications of AI are manifestly “narrow” in nature.

Currently, most software underlying simple applications of AI is still programmed by humans and follows a rule-based approach. This programming approach implies that a programmer defines, first, the problems to be solved by the software and, secondly, how exactly they are to be solved (Russel & Norvig, 2009). While this approach is suitable for clearly defined tasks in a context of few unknown variables that can be mathematically described, such programming is limited by the constrained ability of humans to model highly complex environments (Boulanin, 2016a). In this regard, warfare, which is notoriously fraught with complexity and nonlinearity, is a particularly challenging context. Machine learning provides a possible solution to these limitations, in that it teaches machines how to adapt to new “experiences” of their own accord, instead of programming them to “think” in certain pre-specified ways (Goodfellow, Bengio, & Courville, 2016).

Deep learning, a subtype of machine-learning that is based on artificial neural networks modeled after the human brain (LeCun, Bengio, & Hinton, 2015), is seen as one of the major technological revolutions that could be decisive in gaining tactical and, perhaps, strategic advantage on future battlefields (Tucker, 2017). Deep learning would enable machines to navigate and operate in highly complex environments and anticipate and quickly adapt to a variety of changing circumstances. These abilities are key in situations where mere milliseconds in an action–reaction dynamic can make a critical difference and any mistake—e.g. in the selection of, and attacks upon, human targets—can have the severest consequences. In this respect, the speedy analysis of large amounts of data, which far exceeds the abilities of human analysts (Lewis, 2016; Weisberger, 2017) is of the greatest importance. Where it enables correct judgments, this glut of processing power could perhaps even make attacks more discriminatory, thereby minimizing collateral damage in targeted killing campaigns. The machine-learning algorithms that underlie such capabilities are trained through the consumption of large amounts of real-world training data and once deployed, through experience. In addition to massive computational processing capacity to speed up the learning process, as well the availability of large amounts and varieties of data, machine learning also depends on advanced data storage solutions (Metz, 2017; The Royal Society, 2017).
Apart from hardware and data requirements, the verification of machine-learning algorithms constitutes perhaps the most critical challenge in the development of AWS—one that is particularly relevant in the context of a targeted killing “mission set.” Although the data input as well as the output are traceable, the learning process within a neural network is, essentially, a black box. It is therefore extremely difficult to verify what exact capabilities a system develops, or to predict its resulting behavior with great certainty (Roff & Singer, 2016). This problem is aggravated once the system continues to learn on the basis of its own experiential biases, where inputs can no longer be accurately traced (Defense Science Board, 2016, p. 15). If the capabilities of a system cannot be sufficiently verified, it would be all but impossible to ensure that it is capable of complying, for example, with international human rights law away from the battlefield and the principles of international humanitarian law (IHL) on the battlefield (Wagner, 2016). Moreover, a responsibility gap in the case of system failure is inevitable, if neither the manufacturer nor the commander deploying it can be certain of what exactly the system is capable of at any given moment (Matthias, 2004).

While AI and machine learning are the most significant technological developments in creating AWS, they are not sufficient and require additional “enablers.” Sophisticated sensor technology on board of the system is necessary to collect high-quality data, which can then feed the data analysis process and serve as a basis for classifying and possibly attacking targets. Sensors are also required for the navigation of the system in the air, maritime, or land domains. Finally, to attain intended “kinetic”—i.e. weapons—effects and avoid unintended effects, a platform must be matched with appropriate ordnance, whereas an expendable weapon system must be equipped with a suitable warhead. These technologies are not, however, a unique feature of AWS and could also be used as part of other platforms and weapons.

**The political-military drivers of autonomy**

Although no state has openly declared an interest in developing AWS along the lines of our definition (Sauer, 2016), it can be expected that several political and military drivers will push the arc of technology development further in this direction. In its so-called Third Offset Strategy, the U.S. DoD underscored the key role of robotics in ensuring its military superiority vis-à-vis prospective military competitors (Department of Defense, 2016). Although Deputy Secretary of Defense Bob Work has emphasized that the military’s focus will be primarily on human–machine teaming, or “Centaur Warfare” (Pellerin, 2015), a study by the Future of Life Institute found that technologies currently under development “increasingly replace, not augment, human decision-making” (Tucker, 2016).
While the siren song of “distance, lethality and accuracy” has long been a motivator of advances in both manned and unmanned airpower, autonomous weapons offer additional advantages when contrasted with existing UCAVs. First, operating autonomous systems would allow the military to operate increased numbers of systems simultaneously, in different areas, with minimal human involvement at the tactical level (Marchant et al., 2011). The number of soldiers involved a specific mission executed by autonomous aerial weapons would decrease dramatically in comparison to missions with remotely controlled UCAVs, for example with regard to the analysis of battle-space intelligence that is currently still analyzed by ground elements. Secondly, the communications infrastructure required to operate remotely controlled UCAVs is “highly vulnerable to electronic countermeasures or kinetic attacks on the infrastructure which sustains them” (Sparrow, 2016, p. 6). Systems that can provide ISTAR and strike targets autonomously are not dependent on a permanent communications link with a ground crew, which greatly augments their survivability, especially in contested environments. Thirdly, given the compressed decision-cycles of advanced AIs in comparison with a human decision-maker, autonomous gathering and collation of battlefield intelligence, as well as algorithm-based data analysis, maneuver and target selection, could provide crucial tactical and, by extension, strategic advantages (Lewis, 2016; Weisberger, 2017). Finally, an algorithm-based analysis of battlefield intelligence as a basis for targeting decisions could also potentially make strikes more accurate, thereby minimizing collateral damage and the ensuing public backlash, which would be a highly significant development in the context of targeted killings.

**Autonomous weapons and Western airpower doctrine**

The development of autonomous aerial weapons that are able to identify, track, and attack targets independently of a human operator’s tactical inputs will come about as a result of a confluence of advances in civilian information technology with the political motivation to harness those advances to create a military advantage. The technical capabilities and limitations of the systems in question will depend on the degrees of machine autonomy that political decision-makers, military end users, and systems engineers have chosen to incorporate in them, as well as a host of other engineering considerations. How a system with given characteristics is actually employed within the broader context of a military operation is, however, determined by military doctrine—“the practical expression of [military] theory” (Corum, 1997, pp. 2–3)—and by the exact circumstances of the conflict. Like other military technologies before them, autonomous aerial platforms and weapons will not be introduced into a military-theoretical vacuum, nor will they completely change the way in which Western military forces think about aerial warfare.
The stability of doctrinal preferences

Rather than the AI and hardware components described above, it will be the military-organizational “software” laid down in doctrinal precepts that will determine whether or not armed forces will assign autonomous weapons to the targeted killing “mission set.” The widespread use of autonomous weapons to kill named, or closely specified, individuals would not be based solely on the existence of a doctrinal “fit”—that is, the fact that their use is seen as appropriate and potentially effective within the framework of an established paradigm of military operations. It is, however, unlikely that such uses would occur unless military commanders see a legitimate and effective way of integrating autonomous weapons into the military-theoretical paradigm that they are applying to targeted killing operations. And while doctrine is subject to change, it tends to embody a set of cultural parameters that are fairly stable over time. In other words, the ways in which force is preferably used tend to be “culturally regular” (Lewis, 2012, p. 13), or compatible with a military organization’s ingrained, pre-existing beliefs. Such cultural beliefs tend to be “semipermanent” (Snyder, 1977, p. v): They may be subject to a slow evolution over time, but they do not fundamentally change every time a new technology is introduced, or a doctrine updated, no matter how radical the change may seem.

To provide an extreme example, the United States and the Soviet Union both integrated nuclear weapons into their pre-existing military-theoretical belief systems, which changed surprisingly little in the process. The U.S. continued to adhere to a preferred operational paradigm from the conventional-only era—strategic bombing—in its nuclear planning (Freedman, 2003, pp. 21–23; Kaplan, 2015). The Soviet Union intended to use its nuclear weapons side by side with conventional forces in an all-out drive for victory by any means possible, which again was commensurable with what Soviet military commanders already believed (Lee & Staar, 1986, pp. 23–40; Van Oudenaren, 1986). It was the superpowers’ political-strategic belief system that was upended by the “thermonuclear revolution,” not their fundamental views of how to fight and win a major war, if one were to occur (Jervis, 1989). While it is not strictly impossible, it would seem unlikely that—unlike the nuclear revolution and a host of other large-scale military innovations—the transformation of aerial warfare by autonomous weapons would unfold on a doctrinal “blank slate.”

An ingrained preference for leadership targeting

The approach that undergirds current air power doctrine in the states that are at the forefront of “normalized” targeted killing—the United States, the United Kingdom, and Israel—is a blend of several doctrinal traditions that
have been discussed in considerable detail elsewhere (Olsen, 2015; Pape, 1996, pp. 55–86). Targeted killing belongs to one of these rival clusters of ideas, known broadly as leadership targeting or “decapitation” (Pape, 1996, pp. 79–86). Leadership targeting is premised on the idea that airpower is best used against a closely specified elite at the heart of the enemy’s military or political organization. In the words of airpower theorist John Warden, “only at the center can a single input of energy … result in a significant change in the enemy system” (1997, p. 175) as a whole. Since the late 1980s, this preference—which Warden eloquently expressed, but hardly conjured up from thin air even then—has had a major impact on Western airpower thinking. Among other things, it has given rise to an elaborate paradigm of “effect-based operations,” which was centered on “the idea that an enemy organization’s ability to operate as desired is ultimately more important than destruction of the forces it relies on for defense” (Deptula, 2001, p. 11). While the effects-based operations literature proper is now seen as passé by many soldiers and experts alike, leadership targeting principles remain deeply ingrained in the U.S. and U.K. armed forces (Mattis, 2008). The same is true of the Israel Defense Forces, which made problematic experiences with a hastily implemented “effects-based” paradigm during the Second Lebanon War (Kober, 2008, pp. 32–33), but has maintained an active program of targeted killings both before and since (Bergman, 2018).

Like all forms of leadership decapitation, targeted killing programs in the “war on terror” have rested on the optimistic assumption that terrorist and insurgent networks are hierarchical, centralized, and structured enough to allow attackers “to seriously degrade important enemy functions” (Joint Chiefs of Staff, 2013, p. I–8) in a fairly precise and selective manner. Like all forms of decapitation by airpower, these campaigns sought to disrupt terrorist or insurgent organizations, and perhaps change their internal strategic calculus, by killing off their named, or closely specified, leadership cadres (Hardy & Lushenko, 2012; Johnston, 2012; Jordan, 2009; Price, 2012).

The overall experience of Western armed forces with this approach has been decidedly mixed. The sustained nature of these recent campaigns would seem to demonstrate that the direct impact of eliminating named operatives has been limited and fleeting and that pressure must be constantly maintained through additional strikes. In fact, it has even been argued that the organizational impact on terrorist groups may be counterproductive from the attacker’s perspective (Boyle, 2013; Jordan, 2009, p. 723, 2014a, 2014b). The overall picture of limited effectiveness would seem to be corroborated by classified materials that have been leaked into the public domain (The Intercept, 2015).
Targeted killings in future war

Any significant improvement of the technological base on which recent targeted killing campaigns have rested has the potential to fuel their evolution and further use. Even fairly routine technological advances—e.g. in the areas of sensor capabilities or ordnance—could render the initiation or continuation of a campaign more palatable to political decision-makers. In the following section, we look well beyond the effects of such incremental changes, however, and ask whether the confluence of targeted killings and autonomy could break the mold and lead to normalized application of this problematic means of warfare in contexts far beyond the militarized counter-terrorism response of the past two decades. Specifically, we will look at the potential of targeted killing practices to take hold in high-intensity, conventional and lower-intensity, hybrid conflict scenarios.

Targeted killing and future “conventional” conflicts

As we have emphasized above, the leading exponents of “normalized” targeted killing among Western democracies have already embraced doctrinal frameworks that allow for and, indeed, emphasize leadership targeting in a range of warfare settings. The overall trend in targeted killings over the last several decades has clearly been toward expansion. While the great majority of operations were directed against non-state adversaries, it should not be forgotten that the decapitation approach was originally developed with state actors in mind (Pape, 1996, p. 79). The United States had employed leadership targeting against more traditional adversaries long before it applied a fundamentally analogous and commensurate set of ideas against al-Qaeda. The targeting of high leadership cadres, up to and including the head of government, was an integral part of U.S. military strategy in both Operation Desert Storm (1991) and Operation Iraqi Freedom (2003). In fact, in March 2003, a controversial decapitation strike against Saddam Hussein was executed more or less spontaneously, even before the planned commencement of hostilities (Woodward, 2004, pp. 384–294). Nor was the Iraqi dictator the only foreign head of government who was specifically targeted by the U.S. and its allies. Muammar Gaddafi narrowly escaped U.S. air strikes on one of his residences during Operations El Dorado Canyon (1986). Twenty-five years later, during Operation Odyssey Dawn (2011), the Libyan leader had no such luck and was killed by Western-backed rebels in the immediate aftermath of North Atlantic Treaty Organization (NATO) air strikes on his convoy (BBC, 2011). Hence, the U.S. (and some of its NATO allies) have not only conducted targeted killings outside conflict zones, in a normative vacuum that government lawyers have worked hard to fill (with limited success). It has also struck across international borders, against the leadership cadres of sovereign nations.
Similar leadership decapitation concepts are also being explored by advanced, Western-style military powers outside the NATO context. South Korea has publicly stated a leadership targeting doctrine vis-à-vis the nuclear-armed North. The South Korean military is currently bringing its Hyunmoo-3 cruise missile online with this approach in mind (Reuters, 2013; “South Korean plan to kill Kim,” 2016). While the primary purpose of this shift is almost certainly enhanced conventional deterrence (Mearsheimer, 1985, pp. 23–66), it is apparent that the introduction of advanced means of attack can make such offensive doctrinal frameworks politically palatable to second-tier regional powers as well. While it is not clear that the proliferation of autonomous aerial weapons would follow a pattern that is comparable to cruise missiles, a relatively slow spread to the close allies of leading military powers is a plausible scenario. Another actor that has routinely targeted enemy leadership personnel across international boundaries is Israel. Because these attacks, including Operation Wooden Leg (1985) against the PLO headquarters in Tunis and the systematic leveling of Hezbollah’s headquarters in Beirut’s Dahiyeh neighborhood in 2006 (Lambeth, 2011, pp. 154–157), were directed primarily against non-state organizations, they did not infringe on the sovereignty of the states in question in quite the same way as a targeted killing attempt against their political or military leaders. Nonetheless, it is apparent that the boundaries between targeted killings against non-state actors and interstate conflict are far from absolute.

We believe that the prospect of autonomous aerial weapons being employed for targeted killing purposes in future interstate conflicts of a fairly “conventional” nature is far from implausible. In many such conflicts, the targeting priorities of advanced Western air forces would include striking a range of leadership targets, perhaps including named or closely specified individuals. While this would not necessarily amount to a shadowy targeted killing campaign, it is not difficult to imagine, in the course of a conflict between the world’s most advanced militaries, a “normalization” of individualized micro-targeting of key personnel and vital command cells, enabled by highly advanced sensors, data fusion, and discriminate ordnance. Given that the targeting of military leadership in wartime is not considered beyond the pale even under current norms, “big data” collection by ISTAR assets, military intelligence, and human intelligence sources on the ground could lay the groundwork for such a campaign even in peacetime or during a crisis.

Autonomous aerial weapons could have many potential advantages in such scenarios. Most importantly, they would be much more survivable than non-autonomous unmanned systems in environments that feature a technologically capable adversary, who is able to locate and exploit the electronic emissions of the platform’s data links, to which an “in-the-loop” or “on-the-loop” human controller must have periodic recourse, or attack it with weaponry that
requires extremely fast reaction times to defend against. At the same time, autonomous aerial systems would also feature much greater endurance than manned systems. The advantages of autonomy are, therefore, likely to be most pronounced where there is a need to operate deep inside an advanced defensive perimeter for extended periods of time (Erhard & Work, 2008, pp. 147–227; Scharre, 2014, pp. 20–22). In those circumstances, we would expect that autonomous weapons would be a preferred means of executing counter-leadership strikes, including targeted killings.

The prospect of “hybrid” warfare

Russia’s invasion of the Crimean peninsula in 2014, and its enduring engagement in Eastern Ukraine, have given rise to another line of threat perceptions, which could incentivize the exploitation of emerging technological capabilities and move the targeted killing debate further beyond the now-familiar context of the “war on terror.” Hybrid—also known as nonlinear or sub-conventional—warfare has been traditionally associated with asymmetric conflicts, where weaker parties, such as Hezbollah in the 2006 war against Israel, use a peculiar blend of strategies, technologies, and tactics not usually associated with non-state actors to counter the conventional superiority of their adversary (Deep, 2015). Targeted killings have been a prominent feature of some such conflicts (Bachmann, 2013).

Although hybrid warfare is not a new concept, and has been more widely used than the current debate would suggest (Freedman, 2014), it remains contested and no commonly agreed definition exists in the scholarly community. However, common denominators of existing definitions are that “hybrid warfare” blurs the dividing lines between the binary order of war and peace and create an “in-between” gray zone involving a spectrum of violence. Adversaries “exploit the full spectrum of warfare” by blending conventional, irregular and cyber warfare to weaken the stronger party and to avoid attribution and retribution (Pindjak, 2015). Hence, what is challenging about hybrid warfare is not so much the means themselves but how they are recombined in often ruthless and unexpected ways within a given battlespace (Hoffman, 2007). Such disparate combinations of tactics pose a major challenge to traditional militaries, as they require the flexible and rapid adoption of countervailing approaches and objectives.

In the course of the Ukraine crisis, Russia has undermined a normative order based on established international law and created such a gray zone conflict along its external borders with the NATO. By deploying a blend of tactics that are “coercive and aggressive in nature” but “remain below the threshold of conventional military conflict and open interstate war” (Brands, 2016, p. 1), Russia has triggered a difficult quest for appropriate and effective responses to a “hybrid” threat within a fairly traditional,
interstate setting (North Atlantic Treaty Organization, 2016; Zapfe, 2016). It is likely that advanced Western militaries will continue to encounter the use of such methods in the future.

We would argue that in the light of several of the challenges that emerge from hybrid tactics for traditional militaries, the use of autonomous aerial weapon systems could be perceived as attractive in future sub-conventional wars. State actors faced with “hybrid” tactics by an asymmetrical adversary may be tempted into deploying already existing UCAVs for reasons no more complicated than their availability and cost-effectiveness, whether their opponent also employs them or not. We would go beyond this assessment, however, and argue that the use of hybrid warfare tactics in interstate conflict could serve as an additional incentive for states to use targeted killings by means of AWS.

Micro-targeting of leadership personnel in a fuzzy environment, where misinformation is deliberately proliferated, would require a technological base that provides sufficiently precise information to identify appropriate targets and to execute such strikes. In this regard, an intentional information overload created by hybrid methods will create even greater challenges for data analysts than has been the case in recent asymmetric wars. In order to cope with this growing challenge, machine learning algorithms integrated in aerial (and other) platforms could be much quicker at extracting relevant information. Moreover, the complexity of—and alternation between—disparate tactics in hybrid warfare could further augment the importance of speedy decision-making and a capacity to act before the parameters of the next move once again begin to shift, along with the fluid situation on the ground.

The autonomous identification, selection, and attack of targets could also be more discriminatory than strikes with existing UCAVs. Machine-learning algorithms supported by advanced sensors could not only analyze vast amounts of data more accurately and quickly but also “gather information that will enhance the reliability of identification and permit target engagement when the target is relatively isolated” (Schmitt, 2013, p. 11). Thus, in a sub-conventional scenario, targeted killings with autonomous aerial weapons could be seen as a forceful response against vital elements of the enemy’s military organization, without causing excessive collateral damage. Attacking leadership targets in-theatre could also be perceived as a viable and effective alternative to an expansion of the conflict into the heartland of an aggressive state opponent, which might entail even greater escalatory potential.

Moreover, the nature of AWS, which allows militaries to carry out missions without having to put “boots on the ground” and burdening soldiers with the psychological impact of killings through remote control (Dao, 2013), further simplifies the enduring involvement in “gray zone conflicts” for the states deploying them. In addition, unmarked AWS, which can operate without communications link to a ground crew, provide an even
greater level of ambiguity as to the exact nature of such an involvement than existing UCAVs and would complicate the attribution of targeted killings. This modus operandi, which could be employed both by an attacker and by a determined status-quo power, could create a climate of pervasive uncertainty and fear far beyond the front lines, thereby weakening and demoralizing enemy leaders and populations alike—a likely hallmark of future hybrid operations.

Lastly, in the “war on terror,” the use of UCAVs for targeted killings was enabled by the uncontested nature of the airspace involved (Chamayou, 2015). In a sub-conventional scenario, this condition might no longer be a given. Adversaries with the capabilities to launch electronic or kinetic attacks on a detected system in their airspace will do so to inhibit the gathering sensitive information and carrying out of attacks (Sparrow, 2016). Like in the high-intensity conventional scenarios discussed above, the successful use of targeted killings in a sub-conventional scenario could therefore ultimately depend on the deployment of autonomous aerial systems, due to their improved survivability vis-à-vis remotely controlled UCAVs.

**Scope for regulation**

While ongoing regulation efforts on the international level are not necessarily doomed to fail, we expect that they will have only limited impact for the foreseeable future, thus leaving it largely to individual state actors to implement restraints in their respective policies and practices with regard to autonomous targeted killings.

Following an influential campaign by a coalition of civil society organizations, states took up the issue of AWS in 2013 under the United Nations Convention on Certain Conventional Weapons (CCW). The CCW provides a framework to restrict or ban new types of weapons that are considered “to cause unnecessary or unjustifiable suffering to combatants or to affect civilians indiscriminately” (UN Office Geneva, 2017). It is unlikely, however, that in the near future, a concrete treaty solution, whether a ban or regulatory treaty governing the further development and deployment of AWS, will evolve from this process. After three years, the process is still at an early stage and the outcome remains uncertain. So far, the parties to the CCW have not been able to agree on a common definition of what constitutes an AWS and are therefore still in the process of understanding the subject under discussion. Moreover, competing narratives on what regulatory approach would be most suitable have already emerged, even though the official negotiations have yet to begin (Dunn Cavelty, Fischer, & Balzacq, 2016). The consensus principle of the CCW will further complicate an agreement in this forum. Lastly, the “dual-use” character of many technologies relevant to autonomous weapons makes a preventive treaty on AWS difficult, as,
depending on its design, such a document could also affect many avenues of technological development in the commercial sphere.

On the national level, the U.S. was the first country to outline a policy on “Autonomy in Weapon Systems” in DoD Directive 3000.09 in 2012. In essence, the directive permits the development and use of fully AWS for the use of non-lethal force but requires that a human always be “in-the-loop” when decisions involving lethal force are taken (Department of Defense, 2012). Although the directive has been welcomed as the “world’s first moratorium on lethal fully autonomous weapons,” it has also drawn strong criticism (Human Rights Watch, 2012). For example, while the current directive underscores the obligation to verify, test, and validate a system prior to deployment, these traditional procedures may simply not work for self-learning systems, as already discussed in the first section of the article (Roff & Singer, 2016). Moreover, the policy established in the directive can be overridden by high officials and is valid for only 10 years, if not renewed within the first five years (Department of Defense, 2012). Thus, the requirements established by the directive as well as the rapidly advancing technology development will leave it to the American president Donald Trump—and the leaders of other nations—to determine the further development and potential areas of deployment for AWS—including targeted killings.

Military advantage versus operational control

While it may sound somewhat paradoxical at first glance, a second—and potentially rather significant—barrier to a full-scale, widespread deployment of autonomous aerial weapons for targeted killings could be a set of counter-vailing military imperatives. This comes down to a critical question that has not been a subject of an extensive debate so far: What are practical advantages and disadvantages of the operational use of autonomous weapons for military commanders?

In our examination of this question, we proceed from two assumptions. The first assumption is that commanders will seek to leverage the advantages of autonomy vis-à-vis manned and non-autonomous unmanned systems: increased speed of tactical decision-making, better survivability due to faster reaction times, and the absence of vulnerable data links. The second assumption is that military organizations—network-centric ones, in particular—tend to have a strong preference for close operational supervision of frontline forces (Vego, 2004) and that commanders will be reluctant to surrender control over sensitive operations.

In addition, we believe that the military value of autonomy is likely to be greatest where the maintenance of human supervision imposes a serious operational burden that may outweigh the perceived disadvantages of surrendering direct operational control. As outlined above, this burden will be greatest
in conflicts that feature advanced conventional defenses that put a premium on high levels of survivability and endurance. The compelling advantages of autonomy would increase more or less in proportion to the density and capability of the defensive dispositions in question. They would be greatest in a scenario involving a fellow great power with highly advanced military forces (Friedberg, 2014, pp. 73–84).

Meanwhile, the preference for tightly centralized command arrangements is bound to be strongest in those operations that are most politically sensitive or potentially escalatory, and therefore need to be authorized at the highest levels of government. Recent targeted killings of high-ranking terrorist leaders, most prominently Operation Neptune’s Spear against Osama Bin Laden, provide excellent examples of this. As a result, we believe that advanced Western military forces’ “command philosophies” (Sloan, 2012, pp. 246–250)—which are culturally ingrained and therefore “semipermanent,” just like their doctrinal preferences—will militate against the employment of autonomous weapons, which require surrendering human control, in some types of targeted killing scenarios.

The resistance is likely to be greatest in the most sensitive and potentially escalatory scenarios—which would certainly include targeted killings against high-level personnel across international borders that may in some circumstances involve the potential for escalation to the nuclear level. The resistance will probably be lowest in “routine” targeted killing operations against non-state actors that do not involve a serious infringement on state sovereignty. Because the utility of autonomy would be greatest in interstate warfare against another advanced military force and comparatively low in counter-terrorism operations, where survivability and immediate reaction to technologically advanced threats are not a priority, command philosophy is a likely counterweight to the spillover of “normalized” targeted killing into the interstate conventional and hybrid realms.

The future of targeted killings and international (dis)order

In the above, we have outlined why we believe a changing technological base, the doctrinal preferences of Western armed forces, and the increasing prominence of sub-conventional and, perhaps, conventional conflict scenarios are likely to lead to an expansion of targeted killing practices. But how would the expansion of targeted killing practices beyond the “global war on terror,” as charted above, affect the existing international order?

In the introduction to this special issue, Senn and Troy (2017) draw an important distinction between the behavioral and institutional dimensions of the international order, which are in a dialectical relationship. Whereas the first dimension is marked by “a relatively stable and thus recognizable and predictable pattern of agents’ behavior or its outcomes,” the latter includes
a “set of institutions that emerges from the needs of agents within a social system” (p. 181). The effects of events on both dimensions of the international order can, however, only be understood in the interplay of pre-existing structures and institutions. Therefore, an outlook as to how the possible proliferation of targeted killings could affect the international order can only be provided against the background of the normative status quo, which the transformation of targeted killings has begun to reshape. We would argue that this already burdened normative context may not be entirely transformed by the increasing use of targeted killings beyond the “global war on terror,” but that current concerns about the durability of the “old” order will almost certainly be reinforced.

In the context of military counter-terrorism operations, a number of scholars have studied the impact of targeted killing practices on the institutional dimension of the international order and especially the transformation of pre-existing norms on targeted killing. Fisk and Ramos (2013), for example, have argued that the frequent and unconstrained use of targeted killings has already initiated the development of a new norm on the preventive use of force. This newly emerging norm will likely be strengthened by the proliferation of targeted killings into conventional and sub-conventional scenarios and their expanded use by a broader variety of actors. Adding targeted killings to the legitimate “mission set” in interstate conflicts is also likely to further erode the previously existing anti-assassination norm (Thomas, 2005).

The official justification for the use of force by states, especially when using hitherto unconventional means, can play a key role in the process of transforming the international order (Senn & Troy, 2017, p. 176). In this regard, state actors could start to actively refer to the rising norm, or simply to the erosion of the previous one, in order to justify the use of targeted killings across a broad range of strategic contexts. While we believe that the advantages provided by AWS are most pronounced in conflicts between highly sophisticated adversaries, their use for targeted killing purposes in those scenarios would be a matter of great concern.

For example, at this stage of technology development, it is still difficult to foresee whether AWS will ever be able to comply with principles of IHL on the battlefield including proportionality of the use of force and the distinction between combatants and noncombatants. Moreover, it remains unclear whether—and if so how—individual criminal responsibility for breaches of IHL through autonomous systems could be established, given the opaqueness of algorithms that would make the critical decisions (Wagner, 2016). It is therefore possible that their deployment for targeted killing missions in a broad variety of strategic contexts could significantly weaken some of the foundational pillars of IHL and thereby begin to further erode the institutional dimension of the international order.
Moreover, we believe that an expansion of targeted killings as a result of the availability of autonomous aerial weapons could significantly lower the threshold for the initiation and pursuit of long-lasting gray zone conflicts, with the potential for escalation from sub-conventional into large-scale conventional wars. As a consequence of such scenarios, targeted killings could eventually also affect the behavioral dimension of the international order, challenging and perhaps changing some established and relatively predictable patterns of behavior between state actors.

Conclusions

Our main findings in this article have been twofold. First, we expect that the move toward machine autonomy in aerial weapon systems, which is already a prominent trend in military technology, is likely to offer a limited, but significant, set of advantages to actors engaging in targeted killings. As far as advanced Western militaries are concerned, there is a reason to believe that these systems will be integrated into a pre-existing doctrinal framework that emphasize leadership targeting and that has—in many cases, over the past 15–20 years—found its expression in a propensity for targeted killings. It is also likely that the advantages provided by autonomous aerial weapons will be greatest in conflicts that involve capable opponents with considerable expertise in areas including electronic warfare and air defense.

As a result, we believe that the possibility of targeted killing campaigns taking place in either conventional or sub-conventional (“hybrid”) settings involving state actors—or, perhaps, an idiosyncratic mix of state and non-state actors—are real and should not be underestimated. To the extent that this expectation is borne out by future conflicts, machine autonomy will contribute to a further expansion of targeted killings and to their spread beyond military counter-terrorism operations in the non-state context. In other words, we would contend that the spread of autonomous aerial weapons will likely lead to a significant increase in the scope and incidence of targeted killings in the international system. We also argue that the development of regulatory measures that could hinder or contain the development and deployment of autonomous weapons technologies is unlikely in the foreseeable future, leaving it largely to individual state actors to decide how to use them.

At the same time, we believe that the motivations for the use of such systems will have to compete with countervailing political-military motivations, which result from the ingrained command philosophies of Western-style armed forces and continue to emphasize tight hierarchical control over highly sensitive operations. “Autonomous” aerial warfare, in other words, will remain subject to direct military and political control at the operational and strategic levels, even where adherence to international—and
perhaps even domestic—norms is not seen as a priority. As a result, the above-mentioned effects of the introduction of machine autonomy into the equation of “normalized” targeted killings may be mitigated to some extent by the military and political context within which they are embedded. Overall, then, we would expect both a further muddling of the normative framework of the current international order, within which targeted killings would take place, and the persistence of some significant elements of restraint in their future application.

Note

1. We believe that the U.S. approach to targeted killing under the Obama administration was very much in line with this basic military-theoretical paradigm, notwithstanding the fact that civilian National Security Council staff and the Central Intelligence Agency attained an outsized influence over the specific process by which leadership targeting principles were applied in a particular set of circumstances (Carvin, 2012; McNeal, 2014). Early indications point toward a less centralized procedural approach under the Trump administration (Jaffe & DeYoung, 2017).

Acknowledgements

The authors would like to thank Martin Senn and Jodok Troy for their extensive feedback and support in preparing this article. They would also like to thank Martin Zapfe and the anonymous reviewers for their valuable suggestions, which have resulted in a much-improved version of the original manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Michael Carl Haas is a researcher in the Global Security team at the Center for Security Studies at ETH and a doctoral student at the Institute for Security Policy at Kiel University. His interests include modern air and missile power, the proliferation of advanced conventional weapons, and the management of future interstate military crises. His recent, policy-focused writings have been published inter alia in the RUSI Journal, The Diplomat and The National Interest.

Sophie-Charlotte Fischer is a PhD candidate at the Center for Security Studies at ETH. Prior to her PhD studies, she held a Mercator Fellowship on International Affairs. Her research focuses on the strategic implications of emerging military and dual-use technologies, contemporary arms control practices and technology policy. Her recent publications include a co-authored book chapter on the regulation of AWS in the Routledge Handbook of Security Studies.
Reference list


Lambeth, B. S. (2011). *Air operations in Israel’s war against Hezbollah: Learning from Lebanon and getting it right in Gaza*. Santa Monica, CA: RAND.


