

Killer Robots: UK Government Policy on Fully Autonomous Weapons

April 2013

Article 36 is a UK NGO concerned with policy and legal controls to prevent harm from weapons. In March 2012 the organization called for a ban on fully autonomous armed robots. This paper situates such a ban within a broader argument, for a positive obligation in international law for individual attacks to be under meaningful human control. Article 36 staff are part of the International Committee for Robot Arms Control (ICRAC) and the organization is a founder and Steering Committee member of the Campaign to Stop Killer Robots.

Introduction

Whilst considering that unmanned and automated systems will take on a growing role in its military forces,¹ the UK Government has committed in parliament that the operation of weapon systems will always remain under human control. This commitment provides a positive basis for discussing the effective control of autonomous weapons for the future. However, this commitment is brought into question by a consideration of the UK Ministry of Defence's (MoD) Joint Doctrine Note on 'The UK Approach to Unmanned Systems.'² As a result, current UK doctrine is confused and there are a number of areas where policy needs further elaboration if it is not to be so ambiguous as to be meaningless.

This paper calls on the UK Government to:

- Commit to, and elaborate, meaningful human control over individual attacks.
- Strengthen commitment not to develop fully autonomous weapons and systems that could undertake attacks without meaningful human control.
- Recognize that an international treaty is needed to prohibit fully autonomous weapons and to clarify and strengthen legal protection from autonomous weapons more broadly.

Key elements of UK policy

The UK currently deploys remotely piloted drones in combat operations and UK-based companies are developing systems that have a greater degree of autonomy in their flight and operation – the General Atomics 'Reaper'³ is an armed drone currently in use with UK forces, and the Mantis and Taranis systems are under development by UK-based BAE Systems.⁴

¹ UK Ministry of Defence, 2011 *Joint Doctrine Note 2/11, The UK Approach to Unmanned Systems*. HM Government, 2010 *Securing Britain in an Age of Uncertainty: The Strategic Defence and Security Review*, (<http://www.official-documents.gov.uk/>) and Lord Astor of Haver (Parliamentary Under Secretary of State, Defence; Conservative), Hansard Citation: HL Deb, 26 March 2013, c958 - <http://www.theyworkforyou.com/lords/?id=2013-03-26a.958.0>

² UK Ministry of Defence, 2011 *Joint Doctrine Note 2/11, The UK Approach to Unmanned Systems*.

³ See: http://www.ga-asi.com/products/aircraft/predator_b.php. With respect to the Reaper, the UK Government has stated that, "Reaper is the only UK remotely piloted air system (RPAS) and has been armed with precision guided weapons since May 2008. As of 22 October 2012, the UK Reaper RPAS has provided more than 40,000 hours of intelligence, surveillance and reconnaissance to UK and coalition forces in Afghanistan. 293 Hellfire precision guided missiles and 52 laser guided bombs have been deployed using the UK Reaper RPAS. Reaper is not an autonomous system and all weapons employment depends upon commands from the flight crew. The weapons may be released under the command of a pilot who uses Rules of Engagement (ROE) that are no different to those used for manned UK combat aircraft. The targets are always positively identified as legitimate military objectives, and attacks are prosecuted in strict accordance with the Law of Armed Conflict and UK ROE. Every effort is made to ensure the risk of collateral damage, including civilian casualties, is minimised." See Hansard Citation: HC Deb, 1 November 2012, c370W - <http://www.theyworkforyou.com/wrans/?id=2012-11-01c.125777.h&s=autonomous+weapons#g125777.r0>

⁴ See http://www.baesystems.com/product/BAES_019760/mantis?_afLoop=191979140140000 and http://www.baesystems.com/product/BAES_020273/taranis?_afLoop=191954142880000

The elements of policy considered in this paper consist of parliamentary and political statements in conjunction with the Ministry of Defence (MoD) 2011 *Joint Doctrine Note 2/11, The UK Approach to Unmanned Systems*.

Recent government statements in parliament have tended to assert that strikes using the armed drones currently deployed are controlled by an operator who can therefore ensure existing legal rules are followed. Beyond this, Government spokespeople have made a clear commitment to keeping human control of weapons in the future:

“Let us be absolutely clear that the operation of weapons systems will always be under human control.”⁵

However, whilst political statements have appeared to make unambiguous commitments to keeping weapons under human control the MoD’s *Joint Doctrine* is more ambiguous and includes an explicit statement that attacks without human assessment of the target, or a subsequent human authorization to attack, could still be legal:

“... a mission may require an unmanned aircraft to carry out surveillance or monitoring of a given area, looking for a particular target type, before reporting contacts to a supervisor when found. A human-authorized subsequent attack would be no different to that by a manned aircraft and would be fully compliant with [international humanitarian law], provided the human believed that, based on the information available, the attack met [international humanitarian law] requirements and extant [rules of engagement]. From this position, it would be only a small technical step to enable an unmanned aircraft to fire a weapon based solely on its own sensors, or shared information, and without recourse to higher, human authority. Provided it could be shown that the controlling system appropriately assessed the [international humanitarian law] principles (military necessity; humanity; distinction and proportionality) and that [rules of engagement] were satisfied, this would be entirely legal.”⁶

This assessment illustrates a rather naïve conception of international humanitarian law, which is continued in the paragraph that follows:

“In practice, such operations would present a considerable technological challenge and the software testing and certification for such a system would be extremely expensive as well as time consuming. Meeting the requirement for proportionality and distinction would be particularly problematic, as both of these areas are likely to contain elements of ambiguity requiring sophisticated judgement. Such problems are particularly difficult for a machine to solve and would likely require some form of artificial intelligence to be successful.”⁷

The assertion that the rules of proportionality and distinction would provide problems that it would be “difficult for a machine to solve” suggests that such rules are amenable to definitive solution. Yet it is moral agency that these rules require of humans, coupled with the freedom to choose to follow the rules or not, that are the basis for the normative power of the law.⁸ Rather

⁵ Lord Astor of Hever (Parliamentary Under Secretary of State, Defence; Conservative), House of Lords debate, 26 March 2013, http://www.publications.parliament.uk/pa/ld201213/ldhansrd/text/130326-0001.htm#st_14. The full answer noted: “Fully autonomous systems rely on a certain level of artificial intelligence for making high-level decisions from a very complex environmental input, the result of which might not be fully predictable at a very detailed level. However, let us be absolutely clear that the operation of weapons systems will always be under human control.”

⁶ Ministry of Defence (MoD) 2011 *Joint Doctrine Note 2/11, The UK Approach to Unmanned Systems*.

⁷ Ministry of Defence (MoD) 2011 *Joint Doctrine Note 2/11, The UK Approach to Unmanned Systems*.

⁸ These rules consist of an obligation on commanders to balance certain principles that are in opposition to each other – for example, that the civilian harm anticipated in an attack must not be excessive in relation to the military advantage expected. There is no way of explaining where the tipping point in this balance sits – it requires a moral judgment. Legal scholars and state institutions have different interpretations of how the rules should be applied, what factors need to be considered on either side of the balance, and of what orientation should be taken to uncertainty or lack of knowledge in this balancing process. Within this framework of ambiguity it currently falls to individual human beings to make decisions about what behavior is appropriate. In turn, the law is progressively developed by how human beings argue over and further codify its terms, act within it, and subsequently respond and react to those actions. The body of international law is not a static set of clear-cut instructions, but a contested and dynamic framework for controlling and

than being simply a problem for “software testing” and the like, a movement away from that legal basis would mark a fundamental change in the nature of the law.

Whilst the MoD believes the level of “artificial intelligence” needed to “solve” the problems posed by these legal rules will likely be available in “less than 15 years”, it is perhaps more important that they note the following:

“For operating environments with easily distinguished targets in low clutter environments, a degree of autonomous operation is probably achievable now and data from programmes such as Brimstone and ALARM, for example, would have direct read-across. However, this is unlikely to be of much help to unmanned systems that we expect will have to operate in the future cluttered urban and littoral environments on long endurance missions.”⁹

This formulation is important as it raises issues that are significant for how any movement towards greater autonomy in attacks would likely develop – i.e. by asserting that certain ‘indicators’ of a valid target, in a controlled context, make automatic attack acceptable. Whilst the Joint Doctrine says “probably achievable now”, there are already systems in operation that function in this way - notably ship mounted anti-missile systems and certain ‘sensor fuzed’ weapon systems.¹⁰ For these weapons, it is the relationship between the human operator’s understanding the sensor functioning and human operator’s control over the context (the duration and/or location of sensor functioning) that are argued to allow lawful use of the weapons. However, the Joint Doctrine provides no guidance on how the parameters of either target indicators or operational context should be defined so that in conjunction they are certain to ensure meaningful human control of the individual attack. Parliamentary statements regarding such systems, whilst acknowledging their existence, simply assert that they “would be used strictly in adherence with international law.”¹¹

How “strict adherence” with the law can be maintained without having any delineation of what would constitute meaningful human control over such weapons is an open question. Without such delineation, these automated weapon systems risk a slippery slope towards less human control and greater weapon autonomy. It is the robust definition of meaningful human control over individual attacks that should be the central focus of efforts to prohibit fully autonomous weapons.

Commit to, and elaborate, meaningful human control over individual attacks.

Article 36 believes that in all circumstances there needs to be meaningful human control over individual attacks. Whilst the UK has pledged that there will always be ‘human control’ over the operation of weapon systems, this needs to be further elaborated. There is no situation envisaged where fully autonomous weapons would operate without a human setting certain parameters of their mission, yet that alone is clearly not sufficient human control.

Requirements for meaningful human control over individual attacks include, but are not necessarily limited to:

responding to human behavior. Regardless of the “technological challenge,” allowing machines to make judgments within that framework would not only have a bearing on the outcomes of individual attacks but also on the functioning of the framework as a whole - it would shift the law away from being a structure based on the moral decisions of individual humans. Human beings are not, of course, perfect. They have the capacity to choose to violate rules and to commit abuses. However, more broadly it is the repetitive choosing of compliance with the framework of rules that underpins the normative force of those rules – this is the basis of customary law. Beyond the issue of interpreting legal rules, the act of ‘choosing compliance’ only has value in so far as the actors involved are free to choose otherwise. Yet, it seems unlikely that we would argue for the deployment of machines that could freely choose to violate the laws we want them to comply with.

⁹ Ministry of Defence (MoD) 2011 *Joint Doctrine Note 2/11, The UK Approach to Unmanned Systems*.

¹⁰ See for example: http://en.wikipedia.org/wiki/Phalanx_CIWS and http://en.wikipedia.org/wiki/SMART_155

¹¹ Lord Astor of Hever (Parliamentary Under Secretary of State, Defence; Conservative), Hansard Citation: HL Deb, 26 March 2013, c958: <http://www.theyworkforyou.com/lords/?id=2013-03-26a.958.0>

- **Information** – a human operator, and others responsible for attack planning, need to have adequate contextual information on the target area of an attack, information on why any specific object has been suggested as a target for attack, information on mission objectives, and information on the immediate and longer-term weapon effects that will be created from an attack in that context.
- **Action** – initiating the attack should require a positive action by a human operator.
- **Accountability** – those responsible for assessing the information and executing the attack need to be accountable for the outcomes of the attack.

In addition, it is important that this control relates to **individual** attacks. Whilst it is recognized that an individual attack may include a number of specific target objects, human control will cease to be meaningful if an autonomous weapon system is undertaking multiple attacks that require specific timely consideration of the target, context and anticipated effects.

Strengthen commitment not to develop fully autonomous weapons and systems that could undertake attacks without meaningful human control.

The UK government says it has “currently no intention of developing weapons systems that operate without human intervention.”¹² Furthermore, they have said that, “the MoD’s science and technology programme does not fund research into fully autonomous weapon systems, and no planned offensive systems are to have the capability to prosecute targets without involving a human in the command decision-making process.”¹³

Again, these statements are to be welcomed and provide a positive position from which to strengthen consideration of this issue. By stating that, “no planned offensive systems are to have the capability to prosecute targets without involving a human...” it is indicated that the UK has a positive policy barrier in place to the development of such systems (rather than simply noting that none are currently planned).

However, as indicated above, the MoD’s Joint Doctrine Note appears to keep the option of fully autonomous attacks open and is ambiguous about the target indicators and controls over context that would make other automatic weapons systems acceptable. It is important that further elaboration is provided regarding the level of ‘human involvement’ that is necessary. The UK undertakes reviews of the legality of new weapons as they are developed but the reviews are not made public.¹⁴ If the UK considers human involvement in weapon operation to be necessary for a weapon system to pass through this process it would appear necessary to define the nature of that involvement in order to make that determination. If the UK is not able to provide a more detailed elaboration of this, then it is hard to see how the legal review process can be being implemented effectively. Furthermore, whilst the UK has indicated that it is not developing such systems itself, the Government should also indicate to companies involved in the design and production of weapons that fully autonomous weapon systems are incompatible with UK policy and should not be developed.

Stronger barriers could also be established to the operational deployment of any such systems. The Government has been asked if it would commit to reporting to parliament before deploying any autonomous weapon systems in operations. Whilst reiterating that it didn’t have such systems, the Government did not commit to such reporting.¹⁵ Although the UK has asserted that all weapons are used under rules of engagement that comply with international law, those rules of engagement are confidential.¹⁶

¹² Lord Astor of Hever (Parliamentary Under Secretary of State, Defence; Conservative), House of Lords debate, 26 March 2013, http://www.publications.parliament.uk/pa/ld201213/ldhansrd/text/130326-0001.htm#st_14

¹³ Lord Astor of Hever (Parliamentary Under Secretary of State, Defence; Conservative), written answer to question from Lord Harris of Harringey, 7 March 2013, http://www.publications.parliament.uk/pa/ld201213/ldhansrd/text/130307w0001.htm#wa_st_0

¹⁴ See, UK Joint Service Publication 383, 2004, *The Joint Service Manual of the Law of Armed Conflict*, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/27874/JSP3832004Edition.pdf

¹⁵ See Hansard Citation: HC Deb, 10 September 2012, c82W: <http://www.theyworkforyou.com/wrans/?id=2012-09-10c.119120.h&s=autonomous+weapons#g119120.q0>

¹⁶ See Hansard Citation: HC Deb, 13 September 2011, c1153W: <http://www.theyworkforyou.com/wrans/?id=2011-09-13a.70238.h&s=autonomous+weapons#g70238.r0>

Recognize that an international treaty is needed to clarify and strengthen legal protection from fully autonomous weapons.

The UK has so far argued that an international ban on weapons that can “set and engage targets without further intervention from human operators once activated” is unnecessary because existing international humanitarian law is adequate. Foreign Office Minister Alistair Burt stated that, “the Government considers that the existing provisions of international humanitarian law are sufficient to regulate the use of these weapons and therefore we have no plans to call for an international ban. However, we remain firmly committed to their effective control.”¹⁷

Whilst it can be strongly argued that such systems would be incompatible with the nature and obligations of international humanitarian law, existing rules do not explicitly prohibit such weapons nor do they explicitly require that all individual attacks be subject to meaningful human control. Albeit based on a very limited assessment of what the law demands, we have noted already that the UK MoD’s Joint Doctrine Note is happy to assert that existing legal obligations do not rule out fully autonomous attacks.

Given its commitment not to develop or deploy such systems, if the UK considers such weapons to be incompatible with the existing legal rules it should articulate this and support an instrument that makes this explicit. If the UK does not consider fully autonomous weapons to be incompatible with the existing legal rules then this provides a further reason for making the requirement for meaningful human control over individual attacks, and the attendant prohibition on fully autonomous systems that are incompatible with that obligation, explicit in international and national law.

Through its stigmatizing effect, such legal changes would reduce the likelihood of such weapons being developed and deployed even by states that do not join the instrument. The UK has stated that as yet it has had no discussions within NATO or with the United States about the use of fully autonomous weapons.¹⁸ Legal changes at a national level would also provide a basis for preventing UK companies from financing and undertaking the development and sale of systems that are rejected in current UK policy.

Conclusion

The UK’s current political commitments for weapon systems to always be under human control are to be welcomed. By implication, such political commitments permanently rule out the development and deployment of fully autonomous weapons. However, the political commitments to this effect are brought into question by the contents of the MoD’s Joint Doctrine Note on this issue. The details of UK policy need significant further elaboration if it is to be properly assessed. Further detail on the policy, with respect to automated weapon systems as well as automated weapons, would anyway be required for the UK to undertake its internal weapon review processes. Beyond the national policy, the UK should support international efforts to prevent the development and deployment of fully autonomous weapons by others.

¹⁷ Question from Martin Caton MP, Hansard Citation: HC Deb, 6 March 2013, c1021W: “To ask the Secretary of State for Foreign and Commonwealth Affairs if he will call for an international ban on weapons that can set and engage targets without further intervention from human operators once activated.” Response from Alistair Burt, March 2013: “The use of weapons that can set and engage targets without further intervention from human operators once activated is governed by international humanitarian law as enshrined in the Geneva conventions and their additional protocols. The Government is strongly committed to upholding the Geneva conventions and encouraging others to do the same. The Government considers that the existing provisions of international humanitarian law are sufficient to regulate the use of these weapons and therefore we have no plans to call for an international ban. However, we remain firmly committed to their effective control.” See <http://www.theyworkforyou.com/wrans/?id=2013-03-06a.146005.h&s=Martin+Caton+human#g146005.q0>

¹⁸ Lord Astor of Hever (Parliamentary Under Secretary of State, Defence; Conservative), written answer to question from Lord Harris of Harringey, Hansard Citation, Hansard Citation: HL Deb, 6 March 2013, c405W
See: http://www.publications.parliament.uk/pa/ld201213/ldhansrd/text/130306w0001.htm#w_a_st_2