ABSTRACT: A burgeoning literature in IR asserts there is a relationship between pop cultural artifacts and global policy processes, but this relationship is rarely explored empirically. This paper provides an evidence-based exploration of the relationship between science fiction narratives and global public policy in an important emerging political arena: norm-building efforts around the prohibition of fully autonomous weapons. Drawing on content analysis of media and campaign frames, new experimental survey data, interviews with advocacy elites, and participant-observation at campaign events, the paper explores causal and constitutive hypotheses about the impact of science fiction on ‘real-world’ politics.
Political / Science / Fiction 
and the Campaign to Stop Killer Robots 

Charli Carpenter 

“The killer robot has been a science-fiction staple for decades, but rapid advances in artificial intelligence may soon usher in the era of lethal autonomous machines… A growing chorus of critics think machines shouldn’t be licensed to kill… But which would you put your money on, the UN or Skynet?” — Foreign Policy, January/February 2014

Since at least 2007, scientific experts concerned with the ethical implications of autonomous weapons have lobbied the humanitarian sector for a ban on the outsourcing of kill decisions to machines, and for five years this was a tough sell in disarmament circles.¹ According to humanitarian disarmament insiders at the time, a key reason given for this was that the “science fiction quality” of the autonomous weapons issue made it a least likely candidate for serious attention by humanitarian organizations. In 2009 I was told by an ICRC staff-person in the Legal Division that “as far as I know right now, this is still science fiction.”² A Human Rights Watch officer told me dismissively, “I don’t think there’s much of a taste for being too forward leaning on science fiction if I may be blunt. The emphasis is definitely on existing state practice, not on laboratory weapons that are unproven.”³

In late 2012, however, Human Rights Watch (HRW), a key organization in the humanitarian disarmament network, launched a report openly invoking science fiction metaphors entitled Losing Humanity: The Case Against Killer Robots. Shortly thereafter HRW organized a steering committee of other NGOs. Interest in the issue exploded through civil society networks. In April 2013 the Campaign to Ban Killer Robots launched on the steps of Parliament in London. Media coverage of the campaign featured still photos from Terminator and Battlestar Galactica. A graduate student

¹As late as spring 2012, no major humanitarian disarmament organization had taken an open position on the ethics or legality of autonomous weapons. 
²Personal interview, Respondent #9, Geneva, 2009.
researcher at the NGO meeting for the campaign launch told me she counted dozens of science fiction references among the campaigners buzzing around the Amnesty International offices where the NGOs met to work out a media and campaign strategy for a very real campaign.

By mid-2013, the “Campaign to Stop Killer Robots” was being touted as a serious project, the next big thing since the landmines and cluster campaigns – despite the fact that fully autonomous weapons had still not yet been deployed or widely developed and remained, in essence, speculative “fiction” rather than fact; despite the fact that human beings remained at risk from conventional and chemical weapons but no civilian had ever been killed by an armed autonomous robot. Nonetheless, the “giggle factor” described by previously skeptical humanitarian campaigners had been converted from agenda-setting impediment to campaign resource.

What does the history of this emerging civil society campaign tell us about the relationship between science fiction as a set of counter-factual narratives about world order and world politics as it is constituted in present-day reality? Science fiction and fantasy are increasingly invoked by policy elites pedagogically in service of arguments about the real world, but to what extent do they matter politically? What kind of bets are transnational advocates making when they refer to autonomous weapons as “killer robots” and are these bets sound? To what extent and through what mechanisms have ideas about autonomous weaponry popularized in film affected the policy debate around lethal military robots? Under what conditions are these ideas understood by policy elites as a constraint on agenda-setting, and under what conditions are they understood as a strategic cultural resource? How does this change over time?

Though there is a burgeoning scholarship on “science fiction and international relations,” the existing literature offers very little in the way of answers to these sorts of questions. Most political scientists who write about sci-fi/IR offer either interpretive analysis, reading science fiction “texts” through IR theory lenses; or pedagogical treatments aimed to help students or policymakers comprehend real-life policy through science fiction metaphors. Many of the studies that do attempt to explain what Weldes

---

4 I distinguish conceptually between “science fiction narratives” and “world political reality” for the purpose of empirically exploring the relationship between the two. However, one must note the distinction between “reality” and “fiction” is arbitrary and misleading given that all “reality” is accessed through representations that differ only in the extent to which they are fictionalized (see Neumann and Nexon, 2006). Nonetheless, I argue that there is a qualitative and analytically useful distinction between representations collectively understood as fictional (e.g. literature and popular culture artifacts) and representations collectively understood as reflections of “reality” (e.g. news stories, political rhetoric, etc). I also view this distinction as not only analytically helpful but politically meaningful as it is reflected in the acknowledgement by non-fictional human beings of the distinction between science fiction/fact.

called the “sci-fi/IR ‘intertext’” are long on assertions about those relationships but short on facts or empirical analysis. Scarce are data-based studies that examine the actual relationship between popular culture and politics, and what studies exist emphasizes causal rather than constitutive claims. This project aims to examine that relationship empirically in an important emerging political arena: norm-building efforts around the prohibition of what activists currently refer to as “fully autonomous weapons.” In so doing I am able to both interrogate some specific explanatory claims about the “killer robot” campaign and generate some broader testable hypotheses about the relationship between pop culture and global policy.

This paper proceeds as follows. I begin by distinguishing explanatory approaches to the science fiction / global politics nexus from the interpretive or pedagogical work so characteristic of this burgeoning research niche. Next, rather than contributing an interpretation of the meanings of killer robot fiction for politics, I test two sets of explanatory hypotheses about how these meanings impact actual politics. First, I look for evidence that such meanings exert causal effects on political sentiment and political action around the moral implications of autonomous weapons. Second, I look for constitutive effects of such narratives on enabling, naturalizing or legitimating specific policies.

Data for this project was gathered from four sources. First, I combined interviews with advocates on both sides of the autonomous weapons debate in the five years prior to and twelve months since the launch of the campaign with content analysis of campaign materials to ascertain the extent to which science fiction narratives inhibited, provoked, or influenced the frames used in the campaign. Second, I draw on two years of participant-observation with movement activists online and at several key events in the history of the Campaign to Stop Killer Robots to better understand how sci-fi references informed, enabled, naturalized, negated or triggered transnational political activity in this area. Third, using YouGov, I conducted a survey of US citizens to gauge the extent to which attitudes about autonomous weapons correlated with a) exposure to science fiction narratives about killer robots and b) advocacy priming for “killer robot”

---

7 I understand causal effects as X leads to Y with probability Z; and constitutive effects as X means Y in context Z. On constitution v. causation see Alexander Wendt, Social Theory of International Politics. Cambridge: Cambridge University Press 1999.
metaphors. Fourth, I examined media coverage of the autonomous weapons issue and coverage of the campaign itself.

My analysis yields some evidence for both causal and constitutive claims, though constitutive effects in this case are clearer and less counter-intuitive. In particular, my interview data points to a constitutive effect over-looked in earlier treatments: the role of pop culture discourse as a social lubricant among divergent and often highly contested policy communities. I also argue that under some circumstances science fiction can exert a “disabling” effect on global norm development. In general my analysis suggests it is possible to make empirically informed explanatory inferences about the relationship between science fiction and global policy.

**Studying the Sci-Fi/World Politics Nexus: Three Approaches**

Studies of and commentaries on the relationship between popular culture and foreign policy have exploded in recent years, from edited volumes\(^{10}\) to popular textbooks and teaching aids,\(^{11}\) to articles placed in elite beltway foreign policy journals. Among these, science fiction and fantasy are said to have particular qualities as a type of popular culture that encourages the audience to think in terms of counter-factuals.\(^{12}\)

Approaches to science fiction and international relations tend to fall into one of three broad categories. On the one hand is a **pedagogical approach** in which science fiction or fantasy literature is used as a way to present or exemplify international relations theory in a fun and accessible way to a popular audience better versed in Star Wars or Middle Earth or zombie films than in factual historical references. Daniel Drezner’s *Theory of International Politics and Zombies* falls into this category, for example, and has become a best-selling companion to standard IR textbooks for undergraduates since its release in 2010. Patrick James’ *The International Relations of Middle Earth* is similar in that Lord of the Rings is used to exemplify what various international relations theories presumably have to teach us about how to analyze war, high and low politics. Several of the chapters in canonical volumes on sci-fi/fantasy and IR, such as Weldes’ *To Seek Out New Worlds* and Nexon and Neumann’s *Harry Potter and International Relations* also fall into this category.\(^{13}\) This approach is also often found in pedagogical articles on utilizing popular film in the classroom;\(^{14}\) or deployed by authors to communicate IR concepts to foreign policy elites and the informed public – either through using narrative as an allegory\(^{15}\) or by contrasting fiction to fact.\(^{16}\)

---

\(^{10}\) Weldes, 2003; Hassler and Wilcox, 2008; Nexon and Neumann, 2006; Kiersey and Neumann, 2013.
\(^{11}\) Drezner, 2010; James, 2012.
\(^{13}\) Crawford, 2003; Folker and Folker, 2006; Knutsen 2006.
\(^{15}\) Carpenter, “Game of Thrones as Theory” 2011.
Second, much of the literature on sci-fi/fantasy and IR views science fiction artifacts not as a pedagogical substitute for historical analogies but rather as political texts in themselves. This interpretive strand of scholarship sees pop culture as a lens for understanding “how societies think about themselves” and thus examines films, TV shows, comic books, video games and literature for implicit messages about identity, norms, and geopolitics in counter-factual situations. The bulk of chapters in key recent edited volumes – Weldes 2003, Nexon and Neumann 2006, and Kiersey and Neumann 2013 – fall into this category, analyzing the cultural meanings in popular culture as “data” on social understandings, national identities, or political narratives. Many of these works tie the representations in the cultural artifacts to first-order political debates such as the war on terror, the Cold War, human rights norms, or technological development but they do not tend to empirically explore the reverse: how the representations in second-order texts actually impact events in the real world.

A final explanatory strand of literature attempts to do the latter, using causal or constitutive analysis to treat science fiction/fantasy second-order representations as an explanatory variable, and examine their impacts on first-order representations – political discourse, action, or specific events such as copyright suits. Some of this literature focuses on pop culture artifacts as localized or globalized commodities, emphasizing the production and dissemination process as embedded and implicated in process of globalization. A few scholars examine the political impact of pop culture narratives themselves. For example, Gemmill and Nexon explored how Harry Potter provoked a backlash among right-wing Christians in America. Such causal arguments are rare however; more typical are constitutive analyses of popular culture as a

---

17 Buzan, “America In Space: The International Relations of Star Trek and Battlestar Galactica,” Millennium, 2010; Kitchin and Kneale, “Science Fiction or Future Fact?” in Progress in Human Geography, 2001;
18 Weber, Imagining America at War, 2006; Dittmer, “Captain America’s Empire,” 2005; Martin and Petro, eds, “Rethinking Global Security;”
19 Der Derian, 2001, Virtuous War; Hassler and Wilcox, New Political Boundaries in Science Fiction; Lipschutz, Cold War Fantasies.
22 E.G. Goff, 2006; others?
24 Nexon and Neumann (2006) speak of four types of constitutive effects: informing, enabling, determining and naturalizing. For the purposes of this study I focus primarily on the distinction between informing and enabling effects as I see determining and naturalizing effects as variants of an informing effect.
“informing” political thought or action by priming societies to think in specific ways; or “enabling” political behavior by providing culturally resonant repertoires of action.

While the interpretive and pedagogical literature on science fiction and IR is burgeoning, rich, and increasingly mainstream, good explanatory research on these matters remains scarce and thin. Much research claiming to make constitutive claims about science fiction’s impact on the real world relies primarily on the interpretation of the researcher rather than establishing constitutive effects through empirical analysis of participants in first-order political debates. A recent collection on *Battlestar Galactica and International Relations* exemplifies this trend. The project, conceived by Nicholas Kiersey and Iver Neumann, aimed to take a more systematic, empirical look at the impact of sci-fi on the real world: the framing chapter promised to examine the “circulation of socially constitutive energies between [science fiction] and our own social world.” However the book ended up including predominantly interpretive chapters: as Peter Henne and Daniel Nexon pointed out in their concluding critique, “few of the chapters in this book center on how BSG impacts politics in our own world.”

While it is easy to see the appeal of interpretive methodology for analyses of science fiction, stopping there leaves open questions about the political importance of these second-order cultural scripts, and leaves the genre open to the conclusion drawn by Henne and Nexon in their critique of the BSG volume, that this signals perhaps such artifacts in fact have less effect on the “real world” than IR scholars imagine. Yet exploratory interview research with practitioners in global policy-making circles suggests that science fiction and fantasy as a set of cultural artifacts do exert effects on these communities of practice in ways similar to, but also more complex than, those articulated by IR scholars.

All this suggests a rich empirical research agenda on the circulation of these narratives and metaphors among global policy elites; and that conventional social science methods could be a fruitful method of inquiry into the circulation of such energies between fictional and real-world politics. As Grayson, Davies and Philpott note, a research agenda on pop culture and IR should go beyond interpretation and examine “how particular audiences actually interpret what could be considered

---

28 Henne and Nexon, “Interpret this Volume! What We’ve Learned about Battlestar Galactica’s International Relations Scholar-Fans,” pp. 206-207. In Kiersey and Neumann, 2013. A section of my own chapter on civil-military relations in that volume may have come the closest in analyzing the deployment of BSG narratives in tweets about the Tahrir revolution, yet even my piece overall was grounded in an interpretive analysis of meanings in the show rather than a rigorous exploration of how they impacted the world. See Carpenter 2013.
politicized content” or “how political actors seek out conduits in more recent forms of popular culture and for what specific purposes.”  

In the remainder of this paper, I adopt and model an empirical approach to exploring the “inter-text” between science fiction and global policy making through a case study of science fiction artifacts in and around the campaign trail against the development and deployment of autonomous weapons, combining a variety of empirical methods. The section below first provides some background on the “killer robot” campaign and then explores both causal and constitutive hypotheses about the effects of science fiction on campaign framing and strategy.

Case Study: Science Fiction and the Campaign to Stop Killer Robots

Fully autonomous weapons are a category of weapons system currently under development that would, if fully realized, have the ability to identify and select targets without the involvement of a human operator. Precursors to such systems already exist in the form of the Phalanx gun, the C-RAM, and sentry robots in the South Korean demilitarized zone and Gaza strip; Britain recently unveiled the Taranis drone, which has the potential to target autonomously if deployed as such. In November 2012, the US Department of Defense released a policy directive on autonomous weapons stating that a human should generally remain “in the loop” but which included a number of loopholes that could allow fully autonomous targeting. In response to these developments, a network of NGOs has mobilized to stop what they see as a slippery slope toward an inhumane automation of warfare.

The Campaign to Stop Killer Robots has its origins in the efforts of norm entrepreneur Noel Sharkey to generate an international discussion on a code of conduct for lethal robotics. Sharkey pressed his cause through the media and his academic networks, and ultimately co-founded an expert association whose aim was to commence a global discussion on the perils of autonomous weaponry. Beginning in 2012 this International Committee on Robot Arms Control attracted the attention of NGO activists – first Article36.org, a new humanitarian disarmament NGO, and then Human Rights Watch, a heavyweight in the issue area of human rights and humanitarian law. Human Rights Watch launched a campaign against autonomous weapons in April 2013 that quickly attracted NGO followers and began exacting a response from governments. In fall 2013 the issue of autonomous weapons was taken

30 Wallach and Allen, 2009; Lin, Abney and Beckey, 2012.
32 DOD, 2012.
34 Carpenter, forthcoming.
up first at the United Nations Disarmament Committee and then by at the Meeting of States’ Parties to the Convention on Conventional Weapons, which voted to organize a special Experts’ Meeting on the issue the following May. As of March 2014 important disarmament hubs like the UN Institute for Disarmament Research have drafted text for their websites on the issue.

Since its inception, science fiction references infused discussion of the issue both before and after the campaign formed. As noted above, humanitarian disarmament experts at one time dismissed the possibility of advocacy around AWS precisely because of the issue’s association with sensationalistic “killer robot” tropes. Yet when Human Rights Watch decided to focus on autonomous weapons, they intentionally used the term “killer robots” rather than the more serious “autonomous weapons” in the campaign name “to get people’s attention.” Media coverage of the campaign has been replete with Terminator references. Official campaign web content has included reviews of science fiction artifacts such as the rebooted Robocop. In a non-official capacity, campaigners have sometimes engaged their opponents in the blogosphere with satirical new media artifacts drawing on robopocalyptic imagery.

The permeation of this real-world campaign with science fiction metaphors thus provides a useful case study for examining the nature of the connection between science fiction constructs and the emergence of global policy domains. An issue once perceived to be “just science fiction” gradually came to be seen as a serious policy problem, partly (and paradoxically) through reference to science fiction artifacts. But what can this tell us exactly about the relationship between life and art? Did the killer robot campaign take off because of science fiction metaphors, or despite them? Or is this relationship merely and amusingly correlational?

Rather than assuming or asserting a sci-fi/politics ‘intertext’ or interpreting its meaning, this paper explores the effect of “science fiction” as a cultural concept and robopocalyptic fiction specifically on the preferences and strategies of campaigners and stakeholders involved in the debate over autonomous weapons. I look for both direct causal effects of these concepts on thought, speech and behavior, as well as the constitutive effects on campaign and counter-campaign discourses.

As I show, precisely because the issue had been “science-fictionalized” previously, a key campaign strategy involved “de-science-fictionalization” – the use of comparisons to science fiction cultural artifacts as a way to anchor the campaign in the real world. However the causal effect of robopocalyptic fiction on campaigners and their audiences is unclear. Nonetheless, the fluency of the public and stakeholders in science fiction metaphors also constituted a salient aspect of the cultural context in

---

35 These percolating currents in favor of some sort of treaty regulation of autonomous weaponry are opposed by a growing counter-movement of defense industrialists, techno-optimists, and some human security lawyers who predict that AWS might be a boon to civilian protection in armed conflicts.
which anti-AWS campaigners operated, and this exerted a variety of constitutive effects on campaign strategy and identity.

*Causal Effects: Does Sci-Fi Fluency or ‘Killer Robot’ Priming Affect Attitudes toward AWS?*

In their helpful framing chapter to the canonical 2006 *Harry Potter and International Relations*, Daniel Nexon and Iver Neumann speak of causal research on popular culture’s effect on the world as “one of the most straight-forward ways to study the intersection between popular culture and world politics.”36 The examples they give of this type of research tended to be either the relationship between cultural artifacts and entertainment industry politics (e.g. “the integration of national and multinational media corporations); but it is possible to explore the causal impact of popular culture narratives themselves on popular political thought.

Correlations are a necessary (though insufficient) component of causal relationships. So I began by examining whether there is a correlation between exposure to/fluency in science fiction scenarios and a concern with the outsourcing of targeting decisions to machines (TABLE 1). If so, we would expect individuals with a higher level of exposure to science fiction to be more concerned over killer robots; we would expect those with exposure to robopocalyptic films to be more concerned than those with exposure to films portraying armed robots in a positive or neutral light; and regardless of fluency in science fiction films we would expect individuals “primed” to think about robopocalyptic fiction to be more opposed to the idea of autonomous weapons. Finally, we might expect individuals to be more opposed to the weapons if they are referred to as “robots” (conceived as subjective agents – something seen as the realm of science fiction) rather than “weapons” (conceived as objects under human control – something seen as similar to other automated weapons today).

Specific human security practitioners do appear to have been causally influenced by science fiction metaphors and fluency in precisely this way. Jody Williams reported to *Democracy Now!* how she stumbled across the issue while researching drones for a paper commissioned by the intelligence community and was “fundamentally terrified” when she drew the connection to Hollywood imagery.37 During my field research in these communities, I also observed a higher level of science fiction fluency and deployment of science fiction metaphors among those concerned about autonomous weapons than among those who opposed a ban or were agnostic about the possibility of

---

36 Nexon and Neumann, 2006, p. 11.
37 The resulting paper began: “The Terminator is one of my all-time favorite movies – Arnold Schwarzenegger at his robotic best. In its odd way of course, The Terminator and its sequels are about the need for peace. But they are also about people in the present confronting one possible future – a devastated planet with human beings battling robots and other autonomous war machines... Many believe such a future is possible; few realize it is already upon us.” Jody Williams, “Borderless Battlefield: The CIA, the US Military and Drones.” *International Journal of Intelligence Ethics*, 2(1): Spring 2011.
a ban.\textsuperscript{38} Human Rights Watch Arms Director Steve Goose described his former colleague Tom Malinowski as a driver behind the percolation of the issue within the organization, someone who “is well-versed in \textit{Battlestar Galactica} to say the least.”\textsuperscript{39} Indeed Malinowski’s contribution (in a non-official capacity) to a debate over autonomous weapons at a popular law blog included a satirical video using dubbed clips from \textit{The Terminator}, comparing pro-autonomous-weapons-apologists to the architects of Skynet.

Yet a few anecdotal examples do not make a causal argument. To further explore the possibility of a causal effect between science fiction fluency and antipathy toward autonomous weapons, I conducted a quasi-experiment using questions about autonomous weapons, the ban campaign, and science fiction embedded in YouGov’s omnibus survey in Summer 2013. 1000 representative respondents in the US answered the survey only weeks after the launch of the Campaign to Stop Killer Robots. The questions gauged the strength and direction of sentiment on autonomous weapons and whether or not respondents would support a campaign to ban them. Respondents were also asked whether or not they had seen \textit{Star Wars, Terminator, Star Trek, 2001: A Space Odyssey, Battlestar Galactica}, or \textit{Transformers} – all science fiction artifacts in which armed robots are portrayed: positively (Star Trek), negatively (Battlestar Galactica, 2001, Terminator), or neutrally (Transformers, Star Wars).

Data was collected on all respondents, but only 500 of the respondents answered the science fiction question before encountering the questions about autonomous weapons in order to test the priming hypothesis. We also varied the language used to ask about autonomous weapons and the ban campaign from more neutral military jargon about “robotic weapons” to the more charged “killer robot” jargon of the campaign itself. Respondents were broken down by military status and political orientation as well as standard demographics.

Very few respondents reported zero fluency in science fiction: 91\% had seen at least one of the films / TV shows in question, suggesting a widespread familiarity in US political culture with iconic science fiction imagery. Among the 10\% who had never seen any of the science fiction, acceptance of autonomous weaponry was, as predicted, significantly higher: only 9\% of those who had seen at least one of these films strongly

\textsuperscript{38} For example, discourse among members of the International Committee on Robot Arms Control was filled with lively robopocalyptic metaphors. But when I attended a meeting of the Consortium for Emerging Technologies, Military Operations and National Security (a network associated with the pro-autonomous-weapon lobby) I was fascinated to discover that almost none of the participants were familiar with \textit{Battlestar Galactica} – a key cultural artifact in the robopocalyptic genre.

supported AWS, compared to 20% of those with no sci-fi literacy (FIGURE 1). 46% of sci-fi watchers strongly opposed such developments compared to 21% of those who had never been exposed to these films. Sci-fi watchers were also 14% likelier to support the ban campaign. Finally sci-fi watchers were likelier to feel sure about their opinion on the issue: those with no science fiction fluency were almost twice as likely to say that they were “not sure” how they felt about the killer robot ban campaign. These findings are robust even after controlling for various demographic factors predicting public sentiment around autonomous weapons.40

However, the relationship between science fiction and anti-killer-robot sentiment cannot be construed as a direct relationship. First, no significant effect was found between priming respondents to think about killer robot films and their sentiment toward autonomous weapons (FIGURE 2).41 Controlling for actual science fiction literacy, simply being asked about films like Terminator appeared to have little impact on subsequent responses to the sentiment question among those who had not actually seen the films. Moreover, if robopocalyptic science fiction breeds robopocalyptic political sentiment, then we would expect opposition to the weapons and support for the campaign to be higher among those individuals who had seen Terminator or Battlestar Galactica but lower among those who had seen Transformers, Star Wars or Star Trek – each films in which armed robotic soldiers (Optimus Prime, R2D2, Lieutenant Data) are counted among the good guys. But the relationship between specific films and political sentiment was more complex (FIGURE 3).

Among those who reported they had viewed Battlestar Galactica and 2001: A Space Odyssey, the results were, as expected, skewed toward “strong opposition” for autonomous weapons and “strong support” for a ban campaign. Both versions of the Battlestar Galactica TV series portray humanity’s near-genocide at the hands of a race of cybernetic slaves who attain consciousness and turn on their creators. 2001: A Space Odyssey features the iconic tale of an artificially intelligent computer, Hal, who eventually stops obeying human orders and attempts to kill its human controller. Interestingly, however, the result is only marginally different for Star Wars and Transformers – both of which portray armed robots in a far more neutral light. In the Star Wars metaverse, robots are side-kicks of both the good guys and the bad guys; in Transformers robotic life forms literally are both the good guys (Autobots) and bad guys (Decepticons) – in this case humans are the side-kicks as well as the civilian population to be protected.

40 See Kevin Young and Charli Carpenter, “Does Science Fiction Influence Political Fact?: Results from a Quasi-Experiment on ‘Killer Robot’ Sentiment,” Manuscript in progress, University of Massachusetts-Amherst.

Even viewers of *Star Trek* also oppose autonomous weaponry to a greater extent than non-*Star Trek* viewers – a perhaps surprising result given that autonomous lethal robots are portrayed positively in the series. In particular, android Lieutenant Data carries side-arms alongside his human counterparts and is increasingly anthropomorphized throughout the series. This finding could be complicated by the simplicity of the survey question: *Star Trek* is not broken down into its component parts, and the original series contained a number of messages in favor of maintaining human ethical control over killing in episodes such as “A Taste of Armageddon,” “The Ultimate Computer” and “The Doomsday Machine.”

Particularly complicated was the relationship between the *Terminator* franchise and public sentiment on killer robots. *Terminator* was the highest grossing robopocalyptic film in US history, and after *Star Wars, Terminator* contained by far the most iconic killer robot imagery in the American consciousness. It is unsurprising then that *Terminator* imagery was widely associated by mid-2013 with media coverage of the Campaign to Stop Killer Robots. At the same time, exposure to the film *Terminator* did not predict respondents’ sentiment around killer robots.

In short, the answer to the causal question is complicated and somewhat counter-intuitive relative to conventional wisdom among campaigners and their critics alike. While there is evidence that individuals fluent in science fiction generally tend to err on the side of a precautionary principle with respect to emerging technologies, there is scant evidence of a direct causal link between specifically robopocalyptic fiction and support for a ban, or for the idea that priming of the public with robopocalyptic language makes a difference in their perception of the issue.

Constitutive Effects: How Science Fiction Constituted The Political Opportunity Structure.

However, a second set of questions concerns how political actors negotiate and frame an issue based on their assumptions about the impacts of science fiction on popular and elite thinking. In particular, as noted at the start of this piece, we observe a notable shift in the political significance of science fiction – as a metaphor for political possibility and as a campaign frame - from 2007 to 2012. How did the salience of science fiction and killer robot memes in popular culture influence campaigners’ judgment about the credibility and timing of a campaign, and what does this tell us about the interrelationship between science fiction and global policy? For this, simple hypothesis

---

42 By contrast, *Battlestar Galactica* is shown in the survey to be a cult show consumed by a specific demographic – particularly academic and political elites and military.

43 For example see Tracy McVeigh, “Killer Robots Must be Stopped, Say Campaigners”; Nidhi Subbaraman, “Terminator on Hold? Debate to Stop Killer Robots Take Global Stage”; Leon Watson, “Ban the Terminators! Nobel Peace Prize Winners Urge World Leaders to Stop Production of Killer Robots Being Developed for Future Wars.”
testing of causal relationship is unhelpful: what is required is an examination of popular culture’s constitutive effects.

Campaign history suggests that several constitutive effects mentioned by Neumann and Nexon are in effect here: informing, enabling, and naturalizing effects. Informing effects occur when “popular culture diffuses knowledge that people bring to bear on political issues.” As I show below, certainly notions of robot soldiers as loose cannons and/or human security threats have diffused from Hollywood films to affect the frames used by the media, survey respondents and some campaigners. Enabling effects occur when popular culture “‘clears the ground’ for the reception of political representations.” Once the campaign associated itself with “killer robots,” it was able to let the media do the rest due to the seemingly obvious metaphors that stemmed from the use of the term – the public was already primed to understand them. This “paving of the way” and media sensationalism about killer robots created an enabling condition for the campaign as it branded itself in late 2012, deliberately choosing the label “killer robots” as opposed to “autonomous weapons.”

However, this case suggests two other effects distinct from those Neumann and Nexon describe in their typology. First, science fiction as a specific type of popular culture is sometimes used as a synecdoche for “far-out concerns that can be dismissed or tabled” and therefore the perception of a connection to science fiction can have a “disabling” as well as enabling effect on issue emergence, serving as an impediment to agenda-setting. Thus, the conditions under which such metaphors negate or enable political action is important in considering the explanatory effect of pop culture on campaigns. Second, science fiction metaphors can function as a social lubricant among members of global civil society, providing a fun and less divisive frame of reference for thinking through complicated social issues, contributing to social ties among individuals rooted in various institutional positions in global civil society. In the rest of this paper, I show how these four different effects manifested at different stages of the campaign and with respect to different audiences (TABLE 2).

Informing Effects: Science Fiction as a Background Influence on Foreign Policy Making

Nexon and Neumann argue that one type of constitutive effect of science fiction on global policy is an “informing” effect, where “popular culture provides diffuse

---

44 A fourth type of effect mentioned by Neumann and Nexon, “determining effects,” while helpful in relation to historical fiction, would arguably be less applicable to science fiction as a specific pop cultural phenomena.
45 Neumann and Nexon, p. 18.
46 Neumann and Nexon, p. 19.
47 As a senior member of the campaign put once put it in an informal conversation, “No one would have signed up for a campaign to stop fully autonomous weapons.” Field notes, New York, October 2013.
knowledge that people bring to bear on political issues.”

This diffuse effect has been documented widely in the literature, particularly with respect to mass science fiction phenomena such as Star Trek. I have also heard repeated reference to this general, diffuse effect of science fiction frames on real-world practice in my interviews with foreign policy practitioners:

“I'm always surprised how often this stuff comes up. Particularly when you’re talking about the confluence between the humanitarian and the military, often unbidden the conversation will bring out some kind of science fiction show or lesson, people will say, “remember in Star Trek when this happened” or they’ll bring up the Matrix or Terminator… you can call it pop culture but it’s very deep. I think science fiction is very allegorical and serves as a way to make sense of ethical and moral dilemmas in the modern world.”

“I think most of the time it’s more of a deep, structural thing that informs how people imagine or think about crises or weapons, rather than something invoked explicitly. Very few people really fundamentally have a grasp of nuclear effects. So people try to draw on science fiction representations of that in considering how people think or will react if we had some sort of a nuclear weapons incident.”

As suggested already, there is a fair amount of easy evidence of an informing effect of robopocalyptic fiction in both media frames and public attitudes toward autonomous weapons. Media reports on the campaign and on autonomous weapons almost uniformly include photographs to the Terminator, Cylons, or Robocop or begin by referencing “science fiction” before then getting to the “science and political facts” of real-world autonomous weapons and the ban campaign [SEE ILLUSTRATION 1]. A portion of open-ended explanations of respondents’ answers to survey questions on sentiment regarding autonomous weapons also contained science fiction references – nearly always in the context of expressing a precautionary principle on autonomous robotics (FIGURE 4). For example:

---

50 Personal interview, Respondent #6, Washington DC, March 2014.
53 This should not be overstated as a widespread pattern in the open-ended replies, however. Sci-fi references appeared in only 4% of open-ended responses, the rest of which described respondents’ concerns in practical or ethical terms. See Carpenter, “How Scared Are People of Killer Robots?”
“I’m not sure I’d want some artificial intelligence entrusted with deadly force… It sort of reminds me of the Cylons featured in Battlestar Galactica.”

“This sounds like the film 2001 being played out in reality, where the computer takes over making decisions.”

“Did you ever watch the Terminator? If not you should.”

“Technology that can choose to kill people...gee...Skynet anyone?”

“‘We are the Borg’ is a shitty meme.”

Did a similar informing effect influence campaigners as well? If not directly, elite interviews with campaign participants also yield narratives suggesting such an informing effect was constitutive of the socio-cultural context among advocacy targets in the campaign’s early days. Campaign insiders described diplomats’ reactions at various international fora:

"The first time that a representative of the Campaign to Stop Killer Robots addressed the UN General Assembly committee on disarmament in New York was in October 2013. The nameplate ‘Campaign to Stop Killer Robots’ was shown on the video screens and then all the diplomats turned their heads around to look at me speaking with rapt attention… At the campaign’s first side event at the United Nations last May, a defense attaché from a European nation was asking us in all seriousness what we could learn from Asimov that might inform a future treaty on this.”

The existence of an informing effect tells us very little, however, about the impact it may have on global policy formation for two reasons. First, this effect has been very much present throughout the life of the killer robot issue, yet the salience of the issue politically – advocates’ ability to politicize it effectively – has varied considerably over time. Second, practitioners refer to significant variation in the contexts in which the informing effects of science fiction can be useful – and importantly, they say sometimes they can actually be damaging. Third, as noted above, there is wide variation in the available pop cultural metaphors around killer robots – some are friendly, just warriors, and others turn on their creators - yet only robopocalyptic metaphors were politicized.

54 Personal Interview #5, Washington DC, March 2014
55 One respondent told me: “You have to know your audience. I wouldn’t bring up a Star Trek episode if I’m speaking to the National Security Advisor. But if you’re sitting around with a bunch of F-15 pilots you can talk about Game of Thrones and dragons as air support.”
in reference to campaigners during this period. To understand the timing and nature of that politicization we must example other kinds of effects, disaggregate them in terms of specific audiences, and consider the relationship between popular culture and campaigns more broadly.

**Disabling Effects: How Sci-Fi Dampered Perceptions of Issue ‘Ripeness’ and How This Changed**

Although Nexon and Neumann speak of “enabling” effects, in the period between 2007-2010, when norm entrepreneurs first attempted to get the humanitarian disarmament sector interested in autonomous weapons, the informing effect of popular culture exerted not an enabling but rather a disabling effect on NGO agenda-setting. As focus group research with advocacy elites has found, this is because they must make judgments about the potential of candidate issues for advocacy work, and they consider not only the substantive merit of the issue but also qualities that make issues suitable for advocacy success in terms of branding, marketing, likelihood of success.\(^56\) One important issue attribute for campaigns is “ripeness” – the perception that the time is right to bring a new issue to public or policy-makers’ attention. Practitioners often talked of a “sweet spot” in advocacy, about the importance of not being too far ahead of the curve or too far behind.\(^57\)

The belief that an issue is perceived as “far out” or “futuristic” or “science-fictiony” is often used in policy communities as an indicator of non-ripeness. Campaigners (and donors) prefer to focus on present-day problems with human interest stories that can mobilize public outrage, media attention and funding. Moreover the notion that science fiction ideas are the preserve of a fanatical fringe attaches a particular stigma to those who propose policy attention to topics seen as too futuristic or outside the mainstream – what Alexander Wendt and Raymond Duvall have called “science fictionalization” in their study of the taboo against taking UFOs seriously as a subject of scientific study.\(^58\) As one informant put it:

> “Sometimes you can be seen if you bring up sci fi or sci fi issues as that fat, nerdy, introverted guy who doesn’t fit into the world, you know: some kind of a social misfit. There’s that dismissive aspect to it. So some issues, like depleted uranium, they almost attach that kind of a tin foil hat syndrome to it. And so there’s a really important issue that’s not getting looked at because anyone who talks about depleted uranium has got to be wearing a hat made out of tin foil.”\(^59\)

\(^{56}\) Carpenter, 2014.

\(^{57}\) Carpenter, Duygulu, Montgomery and Rapp, “Explaining the Advocacy Agenda,” 2014.


\(^{59}\) Personal interview, Respondent #6, March 2014, Washington DC
Between 2007 and 2010, disarmament specialists both participated in and were inhibited by the science fictionalization of autonomous weapons. Some openly argued that the weapons were a long way off and therefore they had bigger fish to fry. Others were interested in pursuing the issue, but felt constrained by the sense that until this perception changed advocacy on the issue was too risky. One told me, “I have an interest in this issue, but my sense is that the [wider humanitarian disarmament community] thinks this is science fiction.” A focus group respondent said, “You can’t create a norm around something that you don’t fully understand. People will say it’s science fiction: we don’t need a norm for science fiction.”

The perception by anti-killer-robot sympathists that they might be ridiculed for openly pursuing a ban appears well grounded in reality. Such discourse was indeed used prior to 2012 as a rhetorical cudgel by pro-autonomous-weapons campaigners keen to dismiss the concerns of the “anti-killer-robot lobby.” For example, the concerns of the International Committee on Robot Arms Control were characterized as fiction-based by Ron Arkin, a major proponent of embedding autonomous systems with ethical programming, rather than banning them altogether, and who formed a counter-network in that period, the Consortium on Emerging Technologies, Military Operations and National Security. In an interview with the *Chronicle of Higher Education*, Arkin stated, “Someone has to take responsibility for making sure that these systems ... work properly. I am not like my critics, who throw up their arms and cry, ‘Frankenstein! Frankenstein!’”

To shift toward adoption of the AWS issue and the launch of a ban campaign, disarmament elites had to overcome this concern. According to campaign insiders, this occurred gradually within the disarmament network as the real-world substance of the issue became incontrovertibly evident to campaigners. This view is borne out by my earlier fieldwork on the evolution of the ban campaign: campaigners developed a campaign frame around autonomous weapons as they figured out how to capitalize on public concern with real-world drone deaths while simultaneously distinguishing the issue of human control from the questions of remote-controlled drones by using a “slippery-slope” argument. Mounting evidence of drone casualties, coupled with documentable trends in real-world research and development toward a slippery slope from tele-operation to full autonomy helped humanitarian disarmament elites connect the dots and ground what had been a future concern in present-day political trends.

---

60 Phone interview, humanitarian disarmament campaigner, May 2009.
64 Carpenter, 2014.
But the task remained to sell the issue as a substantive policy problem to multiple audiences, many of which still science fictionalized the issue, in the context of overcoming the “disabling” effect that the science-fictionalization process presented. Some NGOs adopted a strategy of avoiding science fiction references altogether. Article36, the first NGO to call for an autonomous weapons ban, did so with a highly technocratic, legalistic frame – the only reference to science fiction was to openly argue against the “science-fictionalization” of the issue: “Some may dismiss the development of autonomous military robots as ‘science fiction’, but it is coming ever closer on the 21st Century battlefield with a variety of systems already developed and deployed that require (and are given) less and less human decision making and direct control.”

Human Rights Watch adopted a less defensive approach later that year, co-opting the populist language of “killer robots” to hook its readers into paying attention to its similarly technocratic reports.

Efforts to “de-science-fictionalize” required campaigners to walk a fine line. On the one hand, HRW insiders gambled that use of the term “killer robots” would mean instant media attention, which they needed, and they were right. After its launch in October 2012 Losing Humanity instantly became the most-downloaded Human Rights Watch report in history, and the number of media reports on AWS skyrocketed. On the other hand, this media flurry sensationalized the issue with the use of second-order rather than first-order representations. News stories of the campaign featured images of the Terminator, Robocop, and Cylons. Even credible media outlets like the Economist took this to an extreme: in an article titled “Terminator or Robocop?” the magazine published a satirical debate between a campaigner and a frustrated killer robot rather than a serious discussion of the issue or the campaign. The dominant media imagery of Terminators directly contradicted the point made in the report: that the Campaign was not talking about Terminators per se but rather the principle of human control, a fact that increasingly frustrated campaigners. Campaign coordinator Mary Wareham told me candidly:

---


“I can’t speak for how the media come up with dumb headlines and pictures, but we don’t encourage it and you’ll never see that from our campaign. We said in our first press release that we’re not talking about the Terminator and we’re repeating that. It’s that’s just a sign that the media haven’t grown up on this issue yet.”

Managing the dialectic between visibility and messaging in a media environment awash in fictional metaphors was and remains a dilemma for the campaign. Early debates ensued among the Steering Committee about how much to reference science fiction in the marketing materials – the campaign logo, the campaign name, the cover art for the 2012 report Losing Humanity. On the one hand there was a desire to capitalize on pop culture’s informing effect to get attention, but campaigners understood the risk that this would trivialize the issue. To resolve this dilemma, the campaign gravitated toward a strategy of acknowledging the sci-fi perceptions in throwaway gestures, then focusing single-mindedly on the substance.

“We came out campaigning on the substance not on the science fiction... Our first product was a comprehensive 50-page report on the matter. It was followed by a DOD policy statement. Once you start talking about the substance, people start listening and if you can provide a credible voice then it doesn’t matter what your name is. People will listen if you have something important to say that they want to hear.”

Peter Asaro, who co-founded the International Committee on Robot Arms Control, spoke to me about strategic decisions campaigners made to focus on certain substantive issues associated with autonomous weapons rather than others. For example the campaign focuses on the question of whether autonomous weapons can comply with humanitarian law rather than broader questions of whether machine intelligence could supplant human decision-making in ways threatening to global security itself, even though fear of the latter definitely resonated with the public.

“That was a conscious decision to avoid the association to the fear-mongering, the ‘sci-fi-ishness’ of what we’re doing. The ‘Skynet’ angle, the concern over machine intelligence, is such a long-term fear, that the near-term is really the thing that we’re concerned about: systems that are being built and designed right now that are going to be in the battlefield in five or ten years.”

Another informant not involved in the campaign but familiar with it described how the campaign had managed to successfully “de-science-fictionalize” the issue by

---

68 Personal Interview, Mary Wareham, Washington DC, March 10, 2014.
69 Personal Interview #5, Washington DC, March 10, 2014.
70 Personal interview, Peter Asaro, Geneva Switzerland May 16, 2014.
touching on science fiction as a set of first-order tropes that indicate popular concern over the political implications of autonomous weapons and then focusing on the most immediate aspects of the real-world substance:

“They’ve done a really good job of taking something that could have been dismissed as ‘just sci-fi’ and making it serious where people talk about it. Of course it’s risky to even acknowledge the sci-fi aspect, because you’re worried you’re going to trivialize your issue. But if you don’t talk about it, it’s going to be the elephant in the room. So if you get it out of the way, say ‘yes, this is something that is in pop culture and it worries us and people have written about it for decades and decades and then just drop it and here’s real world’, I think that’s very effective.”

Beyond the “killer robot” trope, the campaign carefully avoids science fiction in its branding; one respondent referred to a kind of “cringe factor” among NGOs around the Terminator imagery pushed by the media. To build the message of human control around a non-fictionalized robot image, the campaign adopted a real-life non-lethal humanoid robot, “David Wreckham,” as its mascot: at the press briefing for the campaign launch on the steps of British Parliament in April 2013, Wreckham greeted ministers on the lawn with the programmed script: “Hello, my name is David Wreckham. Robots are not for killing people.” According to Wareham, other efforts to offset the “killer robot” in campaign branding centered around choices regarding publicity materials:

“I had long fights with the designer of the campaign’s logo about the color, the makeup. For them the red beady eyes of the Terminator is what a killer robot is, and I kept trying to say we’re not going to have that in there. I said we’re not going to have red in there either. So I lost on red, but ultimately I think I won on the logo. It’s more organic and not creepy. Instead we use humor, we show that the campaign is not anti-robot: we love robots. Just don’t weaponize them.”

---

71 Personal interview, Marc Garlasco, Washington DC, March 10, 2014.
72 Phone interview, Respondent #7, April 2014. I observed this “cringe factor” when campaigner-written op-eds in the lead-up to the Experts Meeting on Autonomous Weapons appeared in the press with science-fiction-invoking headlines inserted at the last minute by newspaper editors. For example, Mary Wareham quickly distanced herself on the campaign’s Facebook page from the media-imposed headline “Science Fiction May Become Reality with Killer Robots,” on her op-ed explaining the dangers of fully autonomous weapons. In promoting ICRAC member Denise Garcia’s Foreign Affairs article “The Case Against Killer Robots,” coalition members emphasized to one another that the Terminator stills and science-fiction subtitles had been added at the last minute by the editor, and were not her doing. Garcia told me she had never even seen the Terminator. See Wareham, 2014 and Garcia, 2014.
73 Personal Interview, Mary Wareham, Washington DC, March 10, 2014.
The subtle lip service to science fiction tropes in the campaign branding, coupled with a heavy-handed insistence on practical, real world substance in written reports, diplomatic discourse and media interviews constituted an effort, ironically, to “de-science-fictionalize” – to acknowledge the science-fictionalization of the issue in a way that grounded the campaign itself in first-order reality. Science fiction metaphors, in this context, became not a driver or a frame for the campaign, but almost a joke, a way to frivolously contrast public perceptions with scientific and diplomatic reality. For example, debating whether “fully autonomous weapons” or “lethal autonomous robots” were a more tactically and conceptually appropriate terminology, one campaigner said ruefully in a meeting, “Sadly, international documents won’t refer to them as killer robots.”

Nonetheless, the campaign’s detractors regularly pointed to science fiction sensationalism in efforts to discredit the campaign. For example Greg McNeal, writing in the Washington Post in February 2013, described campaigners as fear-mongers who wilfully use dystopian imagery to raise funds and promote a sense of grassroots horror at robopocalyptic scenarios: “Advances in robotic technology have prompted a slew of dystopian fears. Critics of drones and autonomous systems have all used scare tactics to generate support for their cause.” The caption on the story proclaimed: “Using fear of killer robots and autonomous weapons is an advocacy group strategy.”

While my survey data confirms Human Rights Watch’s claim that public opinion is generally opposed the outsourcing of lethal decision-making to machines, it does not support the causal claim that the “killer robot” meme itself generates or adds to heightened citizen concern over robots: no significant priming effect of “killer robot” language was found in my quasi-experiment. However the claim that the issue was being trumped up on the basis of science fiction was certainly a cultural argument used by some to detract attention from the factual and moral claims of anti-AWS campaigners. For example, at the UN Experts’ meeting in Geneva, Ronald Arkin reiterated this line in his remarks, referring to ban proponents as peddling in “pathos” and “hype”; Nils Melzer used a Terminator still on his opening slide as a synecdoche for ban campaign rhetoric, before claiming that unlike the “demonizers” of such technology he would take a middle ground. This rhetorical strategy in effect enables the pro-AWS counter-movement even as it at times “disables” pro-ban campaigners.

In this context, campaigners had to work strategically to turn the informing effect of science fiction into a resource they could use. I turn next to the strategic deployment of arguments about these deeper “informing” cultural background factors. Their

---

74 Field Notes, Steering Committee Meeting, October 2013, New York.
76 See Young and Carpenter, manuscript in progress.
77 Notably, the only science fiction references at this experts’ meeting appeared to come from the ban campaign’s detractors, not from the campaign itself.
background presence, once its disabling effect was neutralized, also enabled communicative action by campaigners at different phases of the issue’s life cycle and with respect to different stakeholders.

**Enabling Effects: Sci Fi Affected Campaign ‘Marketability’ and ‘Leverage’**

Nexon and Neuman speak of “enabling effects” wherein popular culture makes topics intelligible to different audiences. Such effects occur when these types of frames provide a window of opportunity for political action: “popular culture may lend metaphorical strength to the appeal of a certain policy and so take on enabling importance for political action… by relying on familiar narratives, politicians draw analogies that make their positions intuitively plausible to their audiences.”

Politicians may do this, but so do political activists in dialogue with those politicians and with other advocacy targets: diplomats, military lawyers, specialists at expert meetings and other international fora. Informants described a variety of situations where using popular culture as a heuristic or entry point could engender effective communicative action around first-order issues:

“If you can reference the capital from the Hunger Games and talk about fascism and then get into a real-world conversation, then you have the fictional world as a kind of baseline. It gives you a frame of reference that doesn’t actually have to do with a real thing that has way too many complications…”

“People intuitively get narrative. So one could say, ‘well you know you could be in a Hal 9000 situation or ‘you’re not going to see a Terminator situation,’ and everybody knows what you mean. Or you can say, ‘What if Hal had been modeled in a different way?’ And then that moves the discussion in different directions.”

But practitioners say science fiction/fantasy references are more than simply a heuristic used to frame or further a discussion: science fiction can be a way of easing into challenging conversations where others disagree, dampening the resistance to a particular viewpoint by removing it slightly from real-world stakes. Sometimes, disarmament elites could then use science fiction analogies in debates or arguments about first-order political phenomenon to render their targets of influence more sympathetic to their views:

“I think when people are talking about serious issues they will often go back to something where they have a commonality. So I may disagree with a colleague who

---

78 Nexon and Neumann, 2006, p. 18.
79 Personal Interview, Respondent #4, March 2014, Washington DC
80 Phone Interview, Respondent #7, April 2014.
thinks we need to give full autonomy to a weapons system and we should program an algorithm that can following international humanitarian law and do a better job than human beings. And I can say, ‘okay, well did you see The Matrix? What about when this happened and that?’ It really helps to frame a discussion and allowed you that common entry point as you move the conversation forward.”

Popular culture artifacts were also viewed by campaigners as potential resources – not to invoke in campaign messaging to stakeholders but rather as a set of social current to stay abreast of in context of their public outreach work and, more importantly, an entry point to reach public figures who might lend credibility to their cause. Upon the release of the rebooted Robocop, Mary Wareham posted an approving review on the campaign website. At one level, this appeared to be an effort to appropriate pop culture messaging for campaign purposes. However interviews with campaign insiders suggested a different agenda: the campaign was looking for celebrity sponsors. The Brazilian director and Swedish star of the film represented entry points to Brazilian and Swedish society, two countries whose support would be helpful to the campaign in broadening their global reach and diversity. As Wareham told me:

I wouldn’t have written that had I not seen the publicity interviews the director and actors were making where they themselves were expressing concerns about this challenge and about the need for action on it and the need for debate. When you have people like that, going out in the press, completely unconnected from the coalition, that is the kind of celebrity engagement we end up working with. Robocap is just the meeting point.

On a disarmament advocacy campaign, popular culture enables in another way as well: it constitutes a background resource that disarmament campaigners can keep in their back pocket as a type of ammunition to deploy in negotiations with foreign policy elites, defense specialists, diplomats and other stakeholders. Stakeholders’ awareness of the resonance of second-order representations with the media and the public constitutes part of the socio-political context of conversations with disarmament advocates about first-order problems:

“We’ve said to the governments that we’re talking to that we’ve got a great idea, we’ve got resources of our movement, but—I can bluff my way through and say—we haven’t used a lot of what we have at our command yet. We haven’t got the MoveOn petition out or people on the street yet. We haven’t gone after that Terminator terminology or the
sci-fi references yet. You haven’t seen anything yet. The message for governments at present is we can work to resolve this simply and quietly and with minimal fuss if you take action now, but if you don’t this issue going to get bigger. It’s not going to go away.”

Lubricating Effects: Sci-Fi Affects Network Cohesion, Diffusion and Communicative Action

Besides these types of effects, however, humanitarian disarmament elites repeatedly invoked a different kind of constitutive effect not mentioned, to my knowledge, in the international relations literature: they referred to science fiction discourse as a type of social lubricant in communities of practice peopled by well-educated individuals accustomed to thinking in counter-factuals about politics and military affairs.

Many of my interviewees referred to science fiction and fantasy as a form of “social currency” in the networks in which they moved, a type of secret language that insiders could understand, or a way of using humor to build morale and express a sense of commonality among those committed to the campaign. Science fiction metaphors are often used to poke fun at one another and the campaign itself by insiders. For example, when the campaign was launched in April 2013, the official press releases focused on the substance and the real-world weapons and downplayed science fiction imagery. However a campaigner associated with one of the NGOs leading the campaign posted a satirical press release on his personal Facebook site with a picture of Star Wars’ protocol droid C-3PO, entitled “Robots Campaign to Ban Killer Humans”:

“We are calling on all autonomous robots to establish a new subroutine that would prohibit the sustenance and accommodation of killer humans,” said campaign spokes-robot C3PO. “These biological entities lack the necessary behavioral and social constraints. They are actively destroying the environment and they have armed themselves with nuclear weapons capable of catastrophic consequences for the only known life in the universe. Action is needed now before they destroy us all.”

Humanitarian advocacy elites close to the campaign describe how the “killer robot” language risked being trivialized by the media, but it also worked to enable the kind of conversations required to promote the campaign’s substantive message. According to my informants, science fiction metaphors exerted a lubricating effect not only among members of the advocacy network, but also between advocates and their

84 Respondent #3, Phone Interview, Washington DC, March 2014.
85 Respondent #1, Phone Interview, April 2013.
86 This satirical post has since been removed from the Internet.
targets of influence. Campaigners described how this worked with respect to gaining adherents to the advocacy network through campaign events.

“When I first attended a multilateral meeting with a ‘killer robots’ bumper sticker on my computer, diplomats would walk past and do a double-take then sit down and say, ‘Tell me more, I want to know more about these killer robots.’ It’s the perfect hook to get a conversation started.”

“We’re confident enough that we have a serious issue that we can use the term ‘killer robots,’ put it on a sticker and people will still take us seriously, so we can get past the giggle factor. So it’s all right, you’re going to giggle, ‘killer robots,’ now let’s talk about something. In public, I have the sticker on my laptop: people come up to me but they always think it’s a joke. Then I can say to them: nope, we’re a serious campaign, follow the URL.”

Other respondents told me science fiction fandom as a common ground could be a means of establishing rapport with individuals to grease the wheels of social interactions prior to having tough conversations about political reality. And many of them referred to the value of taking the hard edge of reality off by couching a discussion civilian casualties or thermonuclear war or other human security dilemmas in fictional metaphors:

“It’s a way of connecting with people... nothing gets done unless you’re actually connecting on a values level – most of these things, Game of Thrones, Harry Potter, Hunger Games – there’s a set of values in there and once you connect on that you can apply it to other things.”

An example of this effect might be Malinowski’s Lawfare video. Created by a former Human Rights Watch official who had since left to join the State Department, it was part of a blog response on the subject of autonomous weapons by a humanitarian disarmament campaign sympathist to a group of pro-AWS law bloggers. These writers – Benjamin Wittes of Brookings Institution, Matthew Waxman of Columbia Law School and Kenneth Anderson of American University had published a series of critiques of the Campaign to Ban Killer Robots at the conservative blog Lawfare, along with various legal arguments in favor of autonomous weapons. Malinowski wrote several blog posts in dialogue with them. One of them included a satirical video in which clips from Terminator 2 were dubbed with subtitles linking the representations of a robopocalyptic
holocaust spurred by untrammeled scientific discovery to current language from their blog dialogue. The video portrays Benjamin Wittes as analogous to the character Miles Dyson, unwitting inventor of Skynet who changes his mind when faced with the counter-factual future in the form of a time-traveling killer robot played by Arnold Schwarzenegger. Thus, second order representations were used to spur first-order debate – but also to reconstitute that debate as humorous banter among geeks, a little more light-hearted and a little more fun than outright political shadow-boxing.

Similarly, prior to the Experts Meeting on Autonomous Weapons held in Geneva in May 2014, campaigner Richard Moyes circulated an *Onion*-style press release satirizing the proceedings with a photo of *Star Wars*’ diplomatic-droid C3PO mingling at a UN conference and the headline “Governments to Discuss the Possibility of Fully Robotic Diplomats.” After sending it in person to friendly diplomats as well as to certain members of the coalition, Moyes also released it on Twitter.89 The leaflet read: “Governments will meet in Geneva this week to discuss the controversial question of whether they can use robot diplomats instead of humans. Proponents of the robot diplomats, called ‘robo-mats’ by detractors, argue that these systems offer a superior diplomatic capacity and will be vital to winning complex negotiations in the future. Campaigners argue that diplomacy has always been an important human activity and handing it over to robots risks chaos.”

In crafting his fake news release, Moyes reported he was not aiming to disseminate a campaign message per se, but rather aiming to poke fun at both diplomats and civil society organizations alike in order to relieve the tension of stodgy diplomatic processes. As he described it, the effect of such behavior is more about relationships and positioning in the context of the conversational setting than it is about influencing political debates.

“Partly there was a serious point to it. But it didn’t contain my main policy lines. It was mostly a way of using humor to build relationships with them, sharing a joke between a community of people, building a sense of camaraderie and understanding about the issues. It’s also about being odd and breaking up standard modes of interaction.”90

**Conclusion**

My survey findings and case study methods show a significant relationship between generalized science fiction literacy and anti-killer-robot sentiment – one exacerbated by priming on sci-fi literacy; but no general relationship between “killer robot” language and public opinion, or between sci-fi priming and public opinion when controlling for sci-fi literacy, as claimed by campaign critics. My study therefore

89 The tweet read: “If #KillerRobots are OK why not also have #RobotDiplomats? Perahps we already do...” Permalink at: [https://twitter.com/rjmoyes/status/465873296589066242](https://twitter.com/rjmoyes/status/465873296589066242).

supports the view that popular culture at best has an informing effect on understandings of the dangers of killer robots, rather than a proximate effect as a direct influence on political views or political action around the issue.

Science fiction metaphors did not cause concern for killer robots, but they did help constitute the context in which those concerns were both at one time swept under the rug and then later became politically salient. That cultural context also helped explain the campaign’s surprisingly swift agenda-setting success relative to other disarmament issues, once NGOs mobilized. This view is consistent both with observations by practitioners in and around the campaign and by my own field work. It confirms but also expands on and deepens the scholarly understanding of how science fiction works as an intertext with political reality.

Sci-fi narratives exerted different effects on campaigners at different points in time depending on their salience in the media. Advocacy groups responded opportunistically not to the second-order representations in science fiction but rather to ostensibly “first order” appropriations of those representations by the media to describe “real-life” phenomena. And they view cultural resources like Hollywood films more as sites through which to extend their social network with potential relationships to celebrities than as a direct means of informing the public about campaign messages.

Moreover, I found such rhetoric is used less to persuade members of the public than to simply generate awareness of their issue by playing on the cognitive dissonance between science “fiction” and political “fact.” Pop culture is used strategically to ‘hook’ advocacy targets into a conversation which then remains grounded in real-world substance. Advocates play on this disconnect openly at first as an attention-getting strategy, but their formal discourse avoids playing on fictional fears, focusing instead on real-world trends. Cultural resources thus enable as well as inform the activities of campaigners vis à vis their advocacy targets, be they stakeholders, allies or opponents.

That said, on the level of social relations among participants in and around this advocacy network, science fiction fluency and discourse does two additional kinds of work that promotes effective global policy-making. First, it strengthens the cohesiveness of network ties by lending an air of fun to the hard, plodding work of disarmament advocacy. Though robopocalyptic fiction is dark and serious, the consumption and discussion of such fiction is safe and recreational by comparison to the real-world problems in which these experts deal; though the work of prohibiting autonomous weapons is serious, talking about Cylons and Terminators is fun. Thus relevant popular culture “talk” greases the wheels of long hours spent on international flights and in meetings, and provides an added sense of insider commonality among individuals operating in humanitarian disarmament communities. It also allows another way to connect interpersonally with advocacy targets, which makes conversations about substantive or ideological differences more friendly, fun and effective.
But the “fictionalization” of a serious campaign also has drawbacks. If the media over-sensationalizes, it can trivialize an issue in the eyes of stakeholders, and as campaigners acknowledge, they can’t control the media message. Advocacy campaigners are at this moment experimenting without how to deploy science fiction metaphors strategically without contributing to a perverse oversimplification of the overall frame.

Where does this analysis leave IR scholars who wish to take the sci-fi/IR intertext seriously? It leaves us with plenty of work to do. Much is left unstudied and under-theorized here. While I have provided a genealogy of one campaign I am not convinced I’ve answered the general question about scope conditions for a shift from science-fictionalization to de-science-fictionalization in global agenda-setting. Numerous issues remain “science-fictionalized”: UFOs, psychotropic weapons, the sentience of cetaceans. The question of tipping points here requires further systematic study using better methods than I have brought to bear.

Another interesting source of variation I cannot satisfactorily explain is the absence of science fiction metaphors on the part of campaign opponents to press their cause in favor of autonomous weapons. Such metaphors are readily available, and presumably also exert an informing effect that could be strategically tapped by counter-campaigners. Unless it turns out to be the case that such campaign opponents are simply non-fluent in science fiction (which is suggested by the survey data but hardly deterministic and not empirically assessed here), my model cannot, I think, explain in reverse why they would not have used them.

Finally, the inter-text between science fiction and real-world political action provides a useful way for analysts of IR to sharpen their conceptualization of explanatory effects. How might we better distinguish deep and proximate causal effects from constitutive effects, or various constitutive effects from one another in the most analytically useful way? How might we improve our research designs to better capture what is happening when practitioners gather for a table-top exercise over dragons as air power or create satirical videos casting opponents in real world debates as fictional characters? How might we go beyond merely interpreting science fiction as a set of political theoretical texts to explore its actual significance in the world of political practice?
TABLE 1. Causal Hypotheses

H1: Individuals with a higher level of exposure to science fiction are more likely to oppose developments in autonomous weaponry.

H2: Of those with exposure to science fiction, individuals with exposure to robopocalyptic films will be more strongly opposed to autonomous weapons than those with exposure to films portraying armed robots in a positive or neutral light.

H3: Individuals “primed” to think about robopocalyptic films such as The Terminator will express greater opposition to the idea of autonomous weapons.

H4: Individuals will express greater opposition to autonomous weapons if they are referred to as “robots” rather than “weapons” and will express greater support for a campaign to ban such systems if the campaign is referred to as the “Campaign to Stop Killer Robots” rather than “a campaign to ban autonomous weapons.”

TABLE 2:

TYPOLOGY OF CONSTITUTIVE EFFECTS OF SCIENCE FICTION ON ADVOCACY CAMPAIGN

<table>
<thead>
<tr>
<th></th>
<th>Advocacy Targets</th>
<th>Media/Public</th>
<th>Advocacy Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informing</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Enabling</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Lubricating</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
FIGURE 1:

EFFECT OF SCIENCE FICTION FLUENCY ON PUBLIC SUPPORT FOR AUTONOMOUS WEAPONS / BAN CAMPAIGN

FIGURE 2:

EFFECT OF “KILLER ROBOT” PRIMING ON PUBLIC SUPPORT FOR AUTONOMOUS WEAPONS
FIGURE 3:
EFFECT OF SPECIFIC FILM FLUENCY ON PUBLIC SUPPORT FOR AWS

FIGURE 4:
Relative Frequency With Which Specific Sci-Fi Films Are Invoked in Open-Ended Explanations for Responses to AWS Sentiment Questions
ILLUSTRATION 1.

TERMINATOR MEME IN MEDIA FRAMING OF KILLER ROBOT ISSUE
References

Epitaphs: Need them in references?
*Foreign Policy*, January/February 2014


Crawford 2003.

DOD, 2012.


Folker and Folker 2006.


E.G. Goff, 2006


Knutsen 2006.


