

Agent-Based Models of Dual-Use Research Restrictions

Elliott Wagner and Jonathan Herington

Scientific research that could cause grave harm, either through accident or intentional malevolence, is known as dual-use research. Recent high-profile cases of dual-use research in the life sciences have led to debate about the extent to which restrictions on the conduct and dissemination of such research may impede scientific progress. We adapt formal models of scientific networks to systematically explore the effects that different regulatory schemes may have on a community's ability to learn about the world. Our results suggest that, contrary to common wisdom, some restrictions on the conduct and dissemination of dual-use research do not inhibit scientific progress and may actually aid communities in learning.

1. Introduction

Scientific research that poses a significant risk of intentional misuse or unintentional harm is called dual-use research. The classic example is nuclear science, where experiments may yield material or information that can aid the development of atomic weapons (National Research Council [2013]). There is growing recognition, however, that dual-use research occurs in a range of fields, including artificial intelligence, robotics, and the life sciences (Miller and Selgelid [2008]; Boulain and Verbruggen [2017]; Brundage *et al.* [2018]). To illustrate, consider the recent debate over Ron Fouchier's research on avian influenza transmission (Herfst *et al.* [2012]). Fouchier deliberately generated a strain of H5N1 bird flu that is transmissible between mammals. He did this in order to discover which mutations may result in a variant of H5N1 that could cause a devastating human pandemic. This research is an example of a broader class of so-called gain of function experiments, where micro-organisms are deliberately engineered to gain functional capacities—such as enhanced transmissibility or virulence—that increase their pathogenicity. Such research carries a range of risks to human health: from the accidental release of the enhanced pathogen, to the misuse of the published research results to create bioweapons.

Policy proposals for managing the risks of dual-use research introduce two types of restrictions. The first type limits who may perform such research. For example, gain-of-function research may be restricted to teams working in high-containment

laboratories with the hope that this will lower the probability of an accidental release of a highly pathogenic organism (National Institutes of Health [2014]; Baltimore *et al.* [2015]). The second type limits who may learn the results of such research. For example, in the case of Fouchier's research, the National Science Advisory Board for Biosecurity initially decided that the specific mutations found to result in respiratory transmissibility between mammals ought not be published (National Science Advisory Board for Biosecurity [2012]). Such limits on the dissemination of dual-use research aim at reducing the risk that malicious actors will misuse the results in order to create novel bioweapons (Lentzos [2008]; Kaiser and Moreno [2012]).

In debates about whether or not such restrictions on scientific conduct and dissemination are warranted, it is usually taken for granted that these restrictions will have a negative impact on scientific progress.¹ For example, when considering restrictions on the dissemination of Fouchier's results, microbiologist Palese ([2012], p. 115) states that 'publishing these experiments without details is akin to censorship and counter to progress, science, and public health'. Bush ([1945]) offers a classic statement of this intuition. He writes that 'A broad dissemination of scientific information upon which further advancements can readily be made furnishes a sounder foundation for our national security than a policy of restriction which would impede our own progress'. The debate over restrictions has therefore focused on whether the epistemic benefits of unrestricted dual-use research outweigh the risks to human flourishing posed by such research. Yet if restrictions on the conduct or dissemination of dual-use research have no significant epistemic cost, then there is no need to engage in the difficult task of balancing epistemic and human costs. A crucial component of the argument against restrictions will have been undermined.

Our aim here is to systematically investigate the intuition that restrictions on dual-use research will negatively impact scientific progress. We leave to the side questions about the risks posed by dual-use research and focus solely on the widely shared intuition that restrictions will impede the advancement of scientific knowledge. We will investigate this intuition through the lens of multi-armed bandit problems, which have been used by formal epistemologists as a simple model of theory choice in scientific communities.² We do not claim that all cases of dual-use research can be adequately modelled as bandit problems, but we argue that some important cases—vaccine development, nuclear ignition research, and the design of drone control systems—can be usefully studied in this framework.

¹ The assumption that restrictions necessarily hinder scientific progress is widespread among both ethicists (Miller and Selgelid [2008]; Buchanan and Kelley [2013]; Evans [2013]; Douglas [2014]; Imperiale and Casadevall [2014]; Selgelid [2016]) and scientists (Palese and Wang [2012]; Casadevall *et al.* [2014]; Lipsitch and Inglesby [2014]; Fouchier [2015]). Both those who favour restrictions and those who are opposed to restrictions take this intuition for granted.

² This approach was pioneered by Bala and Goyal ([1998]) and Zollman ([2007]); see also (Grim [2009]; Zollman [2010]; Alexander [2013]; Mayo-Wilson *et al.* [2013]; Holman and Bruner [2015]; Kummerfeld and Zollman [2016]) for a partial list of work in this area.

In Section 2 we describe a choice between dual-use and safer techniques that scientists face when developing vaccines for select agents. We then introduce multi-arm bandit problems, and suggest that this choice can be modelled by such problems. We propose that restrictions on dual-use conduct and dissemination can be modelled in bandit problems as restrictions on agents and the flow of information between agents. In Section 3 we use computer simulations to explore whether or not these restrictions do in fact impede scientific progress. We show that not only do restrictions not impede progress, but in some situations restrictions may in fact increase the likelihood of successful learning in bandit problems. In Section 4 we investigate whether or not restrictions slow learning. We show that in some cases restrictions may increase the time it takes a network to learn in bandit problems, but this increase is small. Moreover, in most cases restrictions do not slow learning. Section 5 concludes. In short, we find that the widely held intuition that restrictions on dual-use conduct and dissemination will pose obstacles to scientific progress is not vindicated by these models.

2. Select Agent Vaccine Research

Select agents are micro-organisms that potentially pose a grave risk to human, animal, or plant health.³ The most familiar select agents are those that pose a risk of catastrophic outbreaks of human disease (for example, anthrax, Ebola virus, smallpox) but agricultural pathogens on the list (for example, African swine fever, avian influenza virus) also impose serious economic costs upon the developing world. Research into vaccines for these pathogens therefore has the potential to greatly benefit people in both the developed and the developing world.

Broadly speaking, vaccines come in two flavors: those based on a modified live pathogen and those based on an inactivated pathogen (or subunits of pathogens). The development of modified live vaccines involves modifying a live pathogen through traditional or recombinant techniques, so that it no longer causes clinical disease but is still capable of reproducing within a host. While attenuated strains of select agents are often available, modern genomic techniques allow researchers to produce subtly different vaccine candidates as they search for the most immunogenic strain that does not cause clinical disease (Chambers *et al.* [2016]). On the other hand, inactivated vaccine candidates involve the use of pathogens (or parts of pathogens) that are unable to reproduce within a host. In order to develop vaccine candidates that stimulate the host's immune system and confer immunity, researchers must experiment with different combinations of subunits and adjuvants (that is, immune response stimulating material). Because vaccine immunogenicity relies upon a complex set of factors, it is difficult to predict *a priori* which of these approaches will

³ See <www.selectagents.gov>SelectAgentsandToxinsList.html> for a complete list of micro-organisms classified as select agents.

ultimately produce the most efficacious vaccine for a given select agent in a given host. Nonetheless, if one of these methods produces an initially promising vaccine candidate, then researchers may attempt to refine that candidate in various ways to seek to maximize its immunogenicity. In this respect, select agent vaccine researchers often face a choice between refining existing candidates produced using one method, and exploring the space of possible vaccine candidates that could be developed using an alternative method.⁴

Modified live vaccine research, however, raises serious dual-use and biosafety concerns that are not raised by inactivated vaccine research. The modified live vaccine approach requires the manipulation of live, fully pathogenic strains of the select agent, and thus involves the risk of accidental release of a dangerous pathogen. Moreover, the genetic modification of a select agent (even if it is intended to attenuate its pathogenicity) poses the risk of creating novel, more dangerous strains. And the results of these experiments, if broadly disseminated, may be used by malicious actors to weaponize the pathogen (as was feared with the gain-of-function studies on the H5N1 influenza virus). On the other hand, inactivated vaccine techniques do not pose these additional risks because they neither require the manipulation of live pathogens nor do they aim at engineering novel strains.

So, vaccine researchers choose between a dual-use technique and a technique that does not carry substantial dual-use risks. Note that for many select agents it is not yet known which approach is more likely to lead to superior vaccines (Meeusen *et al.* [2007]; Faburay *et al.* [2017]; Manini *et al.* [2017]). And, since modified live virus research does pose substantial dual-use risks, there are policy questions regarding whether or not this research should be restricted. For example, perhaps the research should only be conducted in a small number of highly secure laboratories? Or perhaps the results of this research should not be openly disseminated, but should rather be shared with only a central controlling organization, such as the Centers for Disease Control? By design, these policies restrict the ability of some scientists to freely explore the space of possible vaccine candidates and refine those candidates that appear most promising. It is precisely these kinds of restrictions that scientists worry will impede the production of scientific knowledge.

2.1. Bandit problems

To test this intuition, we extend a standard model of scientific communities that treats scientists as facing a multi-armed bandit problem. In a bandit problem an agent is confronted with a multi-arm slot machine. The agent, who can pull just one arm at a time, desires to maximize her winnings, but she does not know the rates at which the machine's various arms pay out and thus does not know which arm she should pull. She can, however, learn about an arm's payout rate by pulling it and observing

⁴ For a discussion, see (Chambers *et al.* [2016]).

what happens. As the agent pulls arms she is therefore forced to choose between pulling the arm that appears optimal given the evidence seen thus far, and pulling other, seemingly inferior, arms to investigate whether or not her beliefs about those arms' payout rates are accurate.

Bandit problems have been used to model the dilemma faced by designers of clinical trials (Villar *et al.* [2015]). In this application, each clinical intervention under consideration is an arm of the bandit, and each trial participant is a pull on one of the bandit's arms. There is a tension between exploring the effectiveness of each intervention and exploiting the intervention that, given the evidence seen thus far, appears most efficacious.⁵

Zollman ([2007], [2010]) extends this use of bandit problems, so as to model the dynamics of scientific communities. In his setup, there are several agents all confronted with the same two-arm slot machine. Each arm represents a different theory or research methodology. Each pull of an arm represents an application of that theory or methodology to a new situation or in a new experiment. And the payout represents whether or not the application is successful. The agent's goal is to generate as many successful applications as possible. Each agent is myopically rational, meaning that each agent always chooses to pull the arm that she believes has the greatest payout rate. But, in addition to seeing the results of her own arm pulls, each agent also sees the results of all choices made by her neighbours in the social network. With this framework Zollman and collaborators have investigated how the social network's structure impacts the community's ability to learn which arm truly has the best payout rate. A surprising moral of his research is that sparse social networks—that is, networks in which agents do not have many neighbours—can be conducive to accurate learning.⁶

We can now see how select agent vaccine research can be modelled as a two-arm bandit problem. The modified live virus and inactivated vaccine methodologies are represented by different arms on a slot machine, and a pull of an arm represents the use of that technique to generate a new vaccine candidate. The arm's payout represents whether or not there was an advance in vaccine immunogenicity—that is, a new vaccine candidate that produces a greater immunological response in the target animal than the currently best vaccine.⁷ The alternative vaccine methodologies payout at different rates, and those rates are unknown to the researchers. The ultimate

⁵ Some clinical trials may aim at estimating the success probability of each intervention, but the same tension between exploration and exploitation occurs when only a comparative judgement about the efficacy of each treatment is needed. We would like to thank an anonymous referee for encouraging us to clarify this point.

⁶ This moral has been challenged by Rosenstock *et al.* ([2017]). We agree that the 'Zollman effect' is not robust in Zollman's models, but for the purpose of this article that is beside the point. A robust effect demonstrated below is that restrictions on dual-use conduct and dissemination do not impede scientific progress.

⁷ One might judge success in other ways, of course, such as by the vaccine's efficacy at reducing disease in controlled trials, or the effectiveness of the vaccine in reducing disease rates in the field. For the purpose of our models, however, the precise measure of vaccine success is immaterial.

goal for each scientist is to settle on the vaccine methodology most likely to produce strongly immunogenic vaccine candidates for a particular select agent. Or in other words, their aim is to play the arm with the greatest payout rate.

Other areas of dual-use research can also be modelled by bandit problems. For example, nuclear engineers designing ignition systems for nuclear fusion may choose between direct or indirect laser ignition approaches. This is a dual-use problem because indirect ignition research generates information with proliferation potential, while direct ignition research is far less useful for weapons development (Franceschini *et al.* [2013]; National Research Council [2013]). Incremental improvements to indirect ignition systems have little import for attempts to improve the efficacy of direct ignition (and vice versa). Bandit problems thus work as a simple yet reasonable first-pass model of the quandary nuclear engineers find themselves in. Engineers must choose between pursuing the safe approach and the dual-use approach. It's unknown which approach is actually superior, and there is a tension between exploiting the method that appears superior given evidence seen thus far and exploring the alternative ignition system.

Research into the systems for controlling micro-drones, like those that may eventually be used for targeted spraying of pesticides onto crops, provides another example. Engineers choose between refining existing remotely piloted systems—that is, control systems in which the drone is ultimately piloted by a human operator—and developing or improving fully autonomous systems. Incremental improvements to the piloting interface used by a human operator are of little interest to those seeking to improve autonomous systems (and vice versa). So engineers find themselves in a bandit problem. They choose an approach to pursue without knowing which approach will ultimately prove the most fruitful. Moreover, autonomous control systems arguably qualify as dual-use since such technology could aid the development of autonomous weapon systems (Boulanin and Verbruggen [2017]; Brundage *et al.* [2018]).

The last component of our model is the structure of the scientific community in which researchers operate. Researchers do not act in isolation, but rather are part of a social network through which the results of research are disseminated. For example, when a vaccine candidate is generated in one lab, other labs in the research community may learn whether or not it is an advance over the currently best vaccine. The dissemination of results through this social network is represented by connections between the network's agents. So the track record of successes and failures that an agent observes depends not just on that agent's choice of methodology to pursue, but also on the choices made by her neighbours in the social network.

We will model restrictions on the conduct of research using live select agents as restrictions on an agent's ability to pull one of the arms. Some agents may be limited to pursuing the safe research technique, while other agents may be free to choose either technique according to their beliefs about which is most effective. Similarly, we will model restrictions on the dissemination of dual-use research as restrictions

on the connections between agents in the social network. For instance, it may be that the results of pursuing the killed/subunit vaccine methodology are visible to everyone in the community, while results of modified live virus research are disseminated only to a central authority like the World Health Organization or the Centers for Disease Control. With this setup we can investigate whether the widely shared intuition that restrictions on research conduct and dissemination are obstacles to scientific progress.

3. Results

3.1. The baseline model

To be transparent, we base our models upon those found in (Zollman [2010]), with some tweaks designed to capture restrictions on the conduct and dissemination of dual-use research. Consider a social network with some number of agents who face a two-arm bandit problem. All but one of the agents are actively engaged in research. In each step of the learning process these agents each pull an arm and receive a payout. Payouts are either wins or losses. These payouts are then disseminated to each agent's neighbours in the social network.

The final agent is passive; she does not actually pull arms, but she does learn about the arms' payout rates by observing the results obtained by the active agents. We think of this passive agent as a central authority such as the Centers for Disease Control, or as the manufacturing arm of a pharmaceutical corporation that is charged with actually producing mass quantities of the currently best vaccine candidate. In what follows, we judge the success of networks by whether (and when) this central passive agent correctly learns which arm is superior.

A pull on the first arm of the two-arm bandit yields a win with a probability of 0.5. A pull on the second arm yields a win with a probability of p . In other words, the arms each pay out according to Bernoulli distributions. The agents do not know the objective probabilities with which the arms produce wins.

Each agent has beliefs about both arms' win probabilities. These beliefs are given by beta distributions, which are initially uniform (that is, $\alpha = \beta = 1$). When called upon to pull an arm, each agent chooses the arm that she believes has the greatest probability of producing a win.⁸ After observing the outcome of a pull, an agent updates her beliefs by Bayesian conditionalization. The beta distribution is the conjugate prior probability distribution of the Bernoulli distribution. This useful fact means that if one's prior is a beta distribution and one updates on the outcomes of some Bernoulli trials, then one's posterior will also be a beta distribution. In particular, if the prior distribution is given by parameters α and β , then after observing n

⁸ If an agent believes the arms have equal probabilities of producing wins, then that agent flips a fair coin to pick which arm to pull.

trials of which s were wins, one's posterior will be the beta distribution with parameters $\alpha' = \alpha + s$ and $\beta' = \beta + n - s$.

Without restrictions on either conduct or dissemination, every (active) agent is capable of pulling either of the two arms, and the outcome of every pull is seen by every agent. This is the no-restrictions baseline model, in which no results are censored and no research methods are off limits. To investigate the impact that restrictions may have on scientific progress we can compare the performance of this baseline model to alternatives that (i) limit which agents see the payouts yielded by pulls on a particular arm, (ii) limit which agents can pull which arm, or (iii) apply both kinds of limitations to the same arm simultaneously. In the figures and discussions that follow, each datapoint is the result of 100,000 simulations of this system, each run for 1000 steps.⁹

3.2. Restrictions on dissemination

To model restrictions on the dissemination of dual-use research we circulate the payouts produced by the two arms through different network structures. On the basis of a fair coin flip, one arm is selected to represent the dual-use technique, and the other represents the safe, non-dual-use method. If an agent chooses to pull the dual-use arm, then her results are observed only by herself and the passive agent, which is our stand-in for a central authority like the Centers for Disease Control. In the precise language of graph theory, the results of dual-use methods are disseminated through a star network that has the passive agent as its hub.¹⁰ On the other hand, the results of pulls on the safe arm are seen by all agents. In other words, those results are disseminated through the complete graph.

Figure 1 compares the baseline, no restrictions model to this model in which the dissemination of one arm's results is restricted. To create this figure, the number of agents in the network was fixed at nine (eight active agents and one passive central authority), and the probability p with which the second arm produces successes was varied. As this figure shows, there is no value of p for which restrictions on dissemination impede the central authority's ability to learn which of the two arms is optimal. Indeed, there is a range of p for which restrictions on dissemination actually increase the likelihood that the authority will correctly identify the optimal arm. For example, when $p = 0.6$ the central hub correctly learned which arm is best in 97,213 out of 100,000 simulations, while without restrictions the hub successfully learned which arm is best in only 76,448 out of 100,000. This is an example of what Rosenstock *et al.* ([2017]) dubbed the 'Zollman effect': sparse network structures are sometimes conducive to successful learning.

⁹ All source code for these simulations is available at <github.com/eowagner/bandit_problems>.

¹⁰ Because the hub agent does not conduct experiments herself, the eight active agents are functionally isolated when it comes to the dual-use arm—they cannot learn about the outcomes this arm has produced when other agents have pulled it.

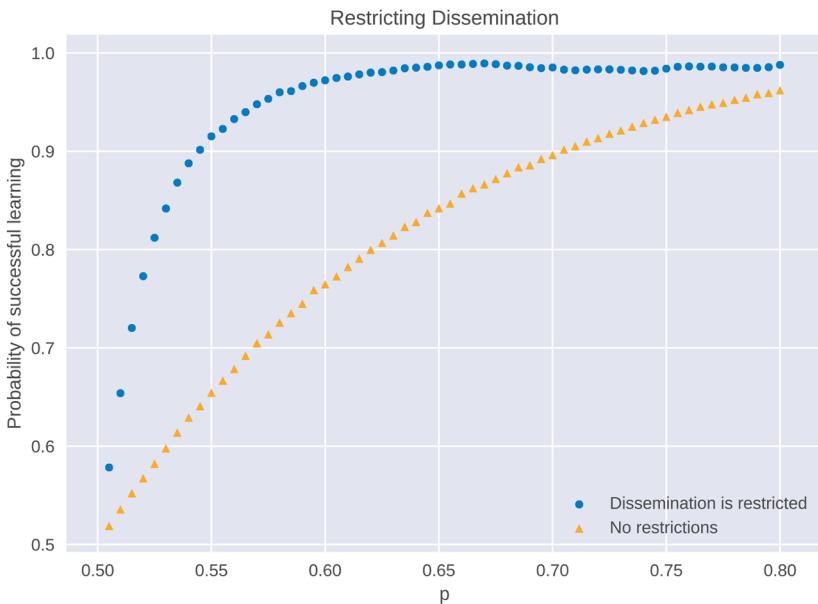


Figure 1. Restrictions on dissemination compared to the no-restrictions baseline model as p is varied. The network has nine agents.

How widespread is the Zollman effect here? As the figure illustrates, the performance of the restriction-less system increases as p increases. This is no surprise since as p grows large, the gap between the payout rate of the first arm (which is fixed at 0.5) and the second arm (which pays out at rate p) also grows. As the difference between arms becomes more glaringly obvious, the epistemic gains provided by a sparse network structure become less pronounced. This corroborates Rosenstock *et al.*'s ([2017]) claim that the Zollman effect is not robust to large differences in the probability of the arms and their observation that as inquiry becomes easier the size of the Zollman effect diminishes.

Why are restrictions on dissemination conducive to successful learning (for at least some values of p)? Zollman's original explanation holds in these models. Restricting the flow of information supports epistemic diversity within the network. In the baseline setup without any restrictions, all agents see all results. This means that it is possible (in fact, it's somewhat likely) that the entire network will be misled by an early biased sample. An early biased sample (for example, one in which initial pulls of the optimal arm produce losses), may lead everyone to mistakenly believe that the objectively better arm is the worse arm. The whole network may then myopically focus on the worse arm, pulling it forever without further exploration of the other, truly better arm. Restricting dissemination makes it less likely that the whole network will be misled in this way because the whole network does not see the results of every pull. In essence, restricting dissemination preserves epistemic diversity (at

least temporarily), leading to more exploration of both arms and thus increased likelihood that the central hub agent will correctly learn which arm is optimal.

Looking again at Figure 1, we notice a surprising fluctuation in the number of successful simulations when p ranges from around 0.68 to around 0.76. This is caused by a phenomenon that only occurs in those simulations in which the objectively worse arm happens to be the arm randomly chosen to have its dissemination restricted. In this case it sometimes happens that many agents choose the worse arm in the first round of learning, and that arm happens to produce several wins. All of these wins are seen by the central hub, who updates her beliefs and ends the round thinking that this arm has a fairly high probability of producing wins. But these results are not disseminated (except to the hub agent), so the agents that pulled this arm will each end the round having seen only one success from that arm and thus end the round with more tempered beliefs about this arm's payout rate. Conversely, pulls of the objectively better arm are seen by everyone, so if just two agents pull that arm and both get wins, then all the active agents will end the round seeing two wins for that arm. So after round one the network as a whole is in an odd situation wherein all active agents agree that the objectively better arm is optimal while the passive hub incorrectly thinks that objectively worse arm is optimal. If p is not too great, this odd situation can persist indefinitely in simulations: every active agent pulls the better arm, but those pulls do not generate enough wins to convince the central hub that this arm is truly best. This phenomenon only occurs when the worse arm is randomly chosen as the arm for which dissemination is restricted, and even then it is quite rare.

Figure 2 shows that this result is stable to changes in the size of the network. As the network grows the performance of the no-restrictions baseline model improves, but even when the network is fairly large, restricting dissemination does increase the likelihood of successful learning. And furthermore, there is no number of agents for which the no-restrictions model outperforms the system with restrictions on dissemination.

Figure 3 illustrates that this result is also robust to changes in the network structure through which the results of dual-use research are disseminated. Real-life proposals for restricting the dissemination of dual-use results often match the star network, that is, a network in which every node is connected to the central hub and no other nodes. But this is a very sparse network, and one might wonder whether our results depend on this exact network structure. They do not. Figure 3 shows the probability of successful learning when dual-use results are disseminated through a handful of alternative network structures. In a k -cycle graph, agents are arranged on a circle, and connected to their k nearest neighbours on both sides. A four-cycle graph with exactly nine nodes is simply the complete graph, in which every agent is connected to every other agent. So the one-, two-, and three-cycles provide increasingly dense dissemination structures for comparison to the no-restrictions baseline in which all agents are connected to all agents. We also considered wheel graphs. A wheel graph is essentially the fusion of a star graph and a one-cycle; in a wheel, there is a hub who

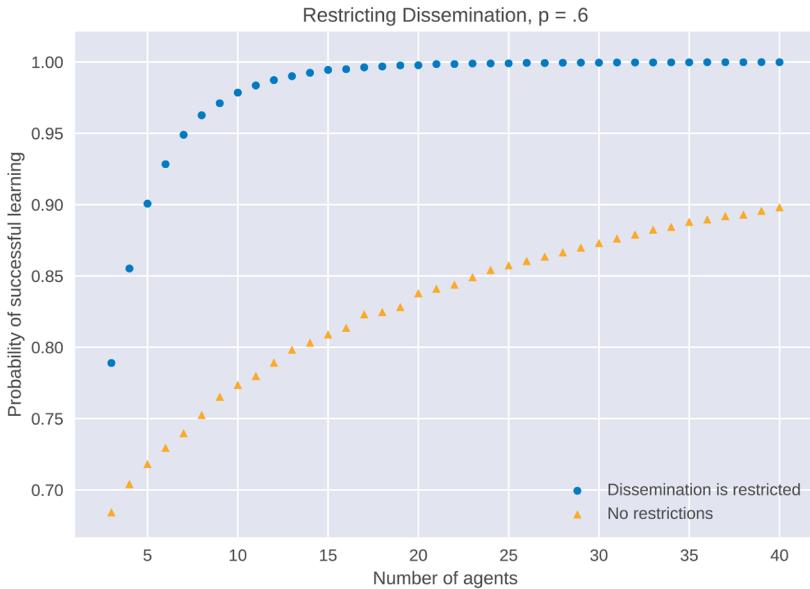


Figure 2. Comparing restricted dissemination to the no-restrictions baseline as the number of agents in the network is varied. p is fixed at 0.6.

is connected to all agents, and all other agents are arranged in a circle on connected to their nearest neighbour on both sides. Figure 3 shows that restricted dissemination does not hinder successful learning no matter the network structure through which dual-use results are disseminated.

So it is a robust fact in these models that restrictions on dissemination do not negatively impact the network's ability to successfully learn which of the two arms is truly optimal. For some parameter combinations— p somewhat close to 0.5 and a low or moderate network size—restricting dissemination actually makes successful

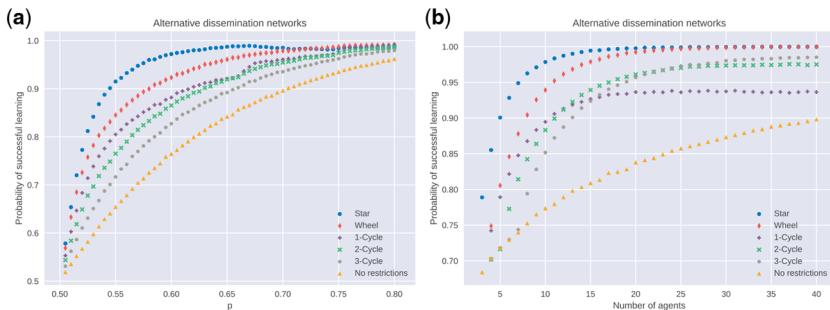


Figure 3. Comparing alternative network structures though which dual-use results are disseminated. On the left the network has nine agents and p is varied. On the right p is fixed at 0.6 and the number of agents in the network is varied.

learning more likely, but there is no combination of parameters for which restricting dissemination inhibits learning.

3.3. Restrictions on conduct

To model restrictions on the conduct of dual-use experiments we limit some agents to just one of the arms in the two-arm bandit problem. Specifically, in each simulation a fair coin is flipped to determine which of the two arms will represent dual-use research, and then some agents are not able to pull that arm. Unlike in previous section, the results of the dual-use arm are not censored, that is, results are disseminated through a complete graph.

Figure 4 compares the baseline, no-restrictions model to models in which between one and seven of the eight active agents are able to pull both both arms (and the remaining agents are limited to just the safe arm). As the figure illustrates, whether or not restricting conduct helps or hinders learning depends on the number of agents who are limited to pulling just one arm. Limiting a small number of agents (between one and four) is conducive to successful learning, but limiting too many agents (six or seven) makes it less likely that the hub will correctly learn which arm is optimal. Limiting five agents is a borderline case that performs slightly better

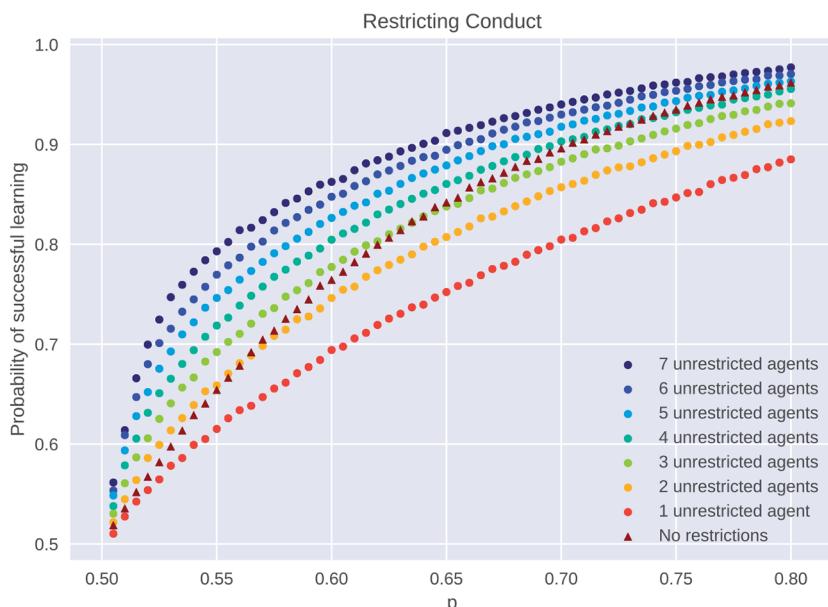


Figure 4. Comparing restrictions on conduct to the no-restrictions baseline model as p is varied. The network has nine agents. An ‘unrestricted agent’ is an agent who is permitted to pull either of the bandit’s arms. The remaining agents are forbidden from choosing the arm that was chosen to represent dual-use research.

than the no restrictions baseline for values of p between 0.5 and 0.6 but slightly worse for values greater than 0.6.

Why is it that restricting conduct is conducive to learning, at least for some values of p and so long as not too many agents are restricted? The explanation here is the same as the explanation for the Zollman effect described above. Preventing some agents from exploring both arms is a way to promote diversity within the network. For example, if some agents are limited to pulling one arm, then it is impossible for the entire community to mistakenly abandon that arm after a misleading sample of payouts. Of course, if too many agents are limited then this diversity is lost. In fact, limiting too many agents has the opposite effect of imposing homogeneity upon the network; the network will be stuck pursuing just one of the arms and unable to effectively explore the other.

Figure 5 shows how this effect varies with the size of the network. As the network grows, the gap between the success rate of the no-restrictions baseline and restrictions on conduct shrinks. But no matter the network's size, restricting just some agents to only one of the two arms does not lower the likelihood of successful learning.

3.4. Combined restrictions on both dissemination and conduct

Lastly, consider combined restrictions on dissemination and conduct. One of the bandit's two arms is randomly chosen to represent dual-use techniques, only some agents

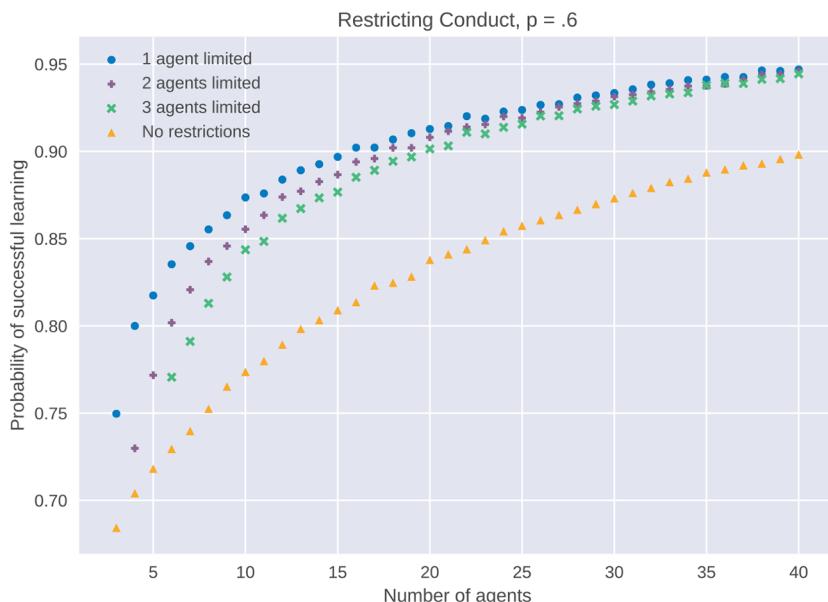


Figure 5. Comparing restrictions on conduct to the no-restrictions baseline model as network size is varied. p is held fixed at 0.6.

are permitted to pull this arm, and the payouts yielded by this arm are disseminated only to the network's central authority. What impact does this combination of restrictions have on the network's ability to identify which of the two arms is optimal?

Figure 6 compares the performance of such combination models to the no-restrictions baseline as p is varied. This figure shows that the combination model outperforms the no-restrictions baseline as long as at least two agents are permitted to pull the dual-use arm. And moreover, as before, for at least some values of p , simultaneous restrictions on both dual-use conduct and the dissemination of dual-use results increase the likelihood that the network will successfully learn which arm is optimal.

Figure 7 shows how the three regulatory schemes compare to each other and to the no-restrictions baseline. When just one agent is limited to the non-dual-use arm, then restricting just dissemination and restricting both dissemination and conduct show nearly identical success rates, and those are the best performing models. Restricting just conduct is worse than restricting both conduct and dissemination simultaneously. And the no-restrictions baseline is worst of all.

So what do we learn? For some parameters, we see a Zollman effect in which restrictions increase the likelihood of successful learning. But more importantly, there are no parameters for which restrictions on dissemination impede successful learning. And, as long as not too many agents are prohibited from investigating both arms

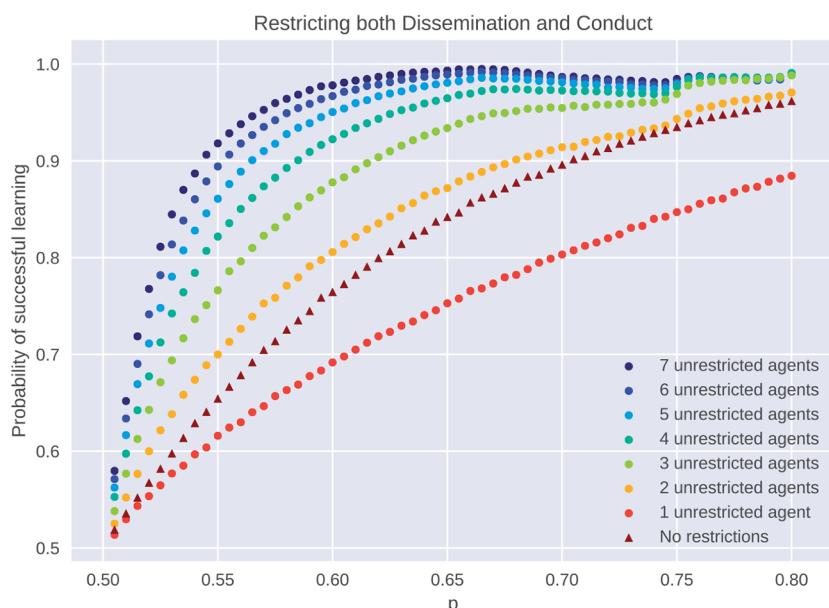


Figure 6. Simultaneous restrictions on dissemination and conduct compared to the no-restrictions baseline model. There are nine agents in the network.

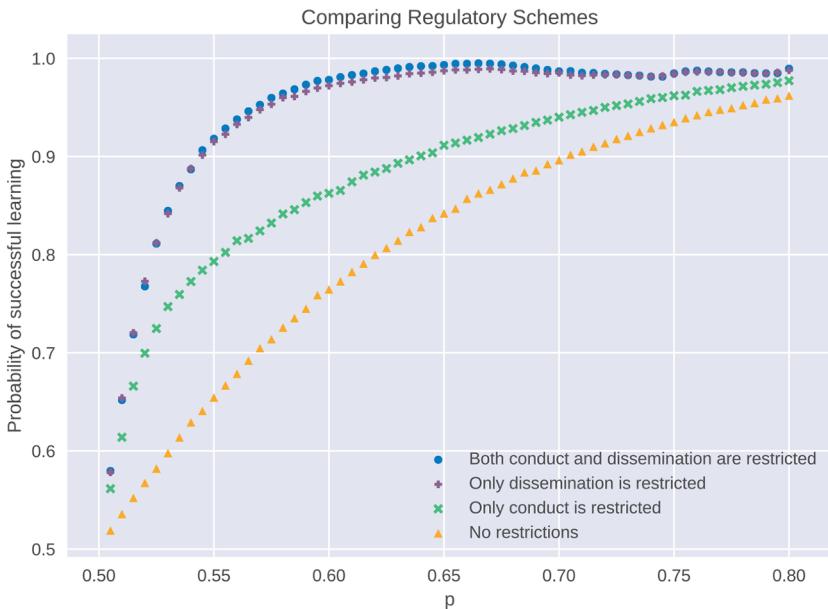


Figure 7. The effects on various regulatory proposals on the likelihood of successful learning.

(for example, with eight agents, that at least two can pull both arms) then there are no values of p for which restrictions decrease the likelihood of successful learning.

4. Do Restrictions Slow Learning?

So, restrictions on conduct and dissemination do not appear to inhibit the network's ability to learn which arm is optimal, but do restrictions slow down the rate at which the network learns? The short answer is yes, but only in some cases and only slightly. To address it in a systematic manner, we will consider the number of steps that it takes the central hub agent to lock on to the optimal arm. In other words, we will look at only those simulations in which the central agent has correctly identified the optimal arm at the simulation's end (that is, after 1000 steps) and count the number of steps before the central hub settles on this arm. We call this the simulation's lock time.

First, let's focus on parameters within the range for which the Zollman effect is seen. Consider $p = 0.6$ and networks with nine agents. Most of these simulations settle on an arm quite quickly. For example, after just two steps of learning, 44,765 out of 100,000 simulations of the no-restrictions baseline model had locked on to the correct arm. And after five steps of learning, an additional 14,834 simulations lock on to the correct arm. The mean number of steps until lock on is approximately 8.3.

Compare this to the model with restrictions on dissemination. In this model, only 35,948 simulations lock on after two steps of learning. But an additional 16,127 lock

on after five steps. The mean number of steps until lock is 18.1, but this average does not tell the whole story. Although the mean number of steps until successful lock is somewhat longer here than in the no-restrictions baseline, the raw number of successful simulations is so much greater that more simulations of the model with restrictions successfully lock on after twelve steps than had simulations of the no-restrictions baseline by the same number of steps.

Figure 8 illustrates this distributional information with a cumulative histogram. The height of each bar shows the total number of simulations (out of 100,000) that had successfully locked on to the optimal arm by the given number of steps. As described above, the no-restrictions baseline performs slightly better after a very small number of steps. But the restricted dissemination model quickly surpasses the baseline model.

This same is true for restrictions on conduct alone and also for simultaneous restrictions on both dissemination and conduct. Figure 9 contains cumulative histograms illustrating this same comparison for those regulatory schemes. As the figures show, more simulations of the baseline model lock on after a very small number of steps. But the models with restrictions quickly catch up and surpass the baseline after around twelve steps.

What about when parameters are such that the Zollman effect is not present? That is, parameters for which restrictions neither help nor harm the likelihood that the network will successfully learn which arm is optimal. In these cases restrictions do not

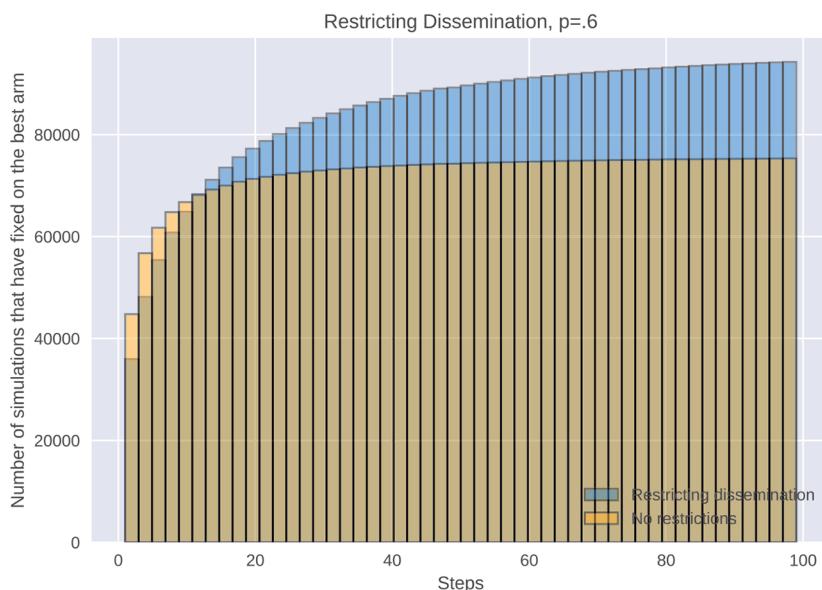


Figure 8. Cumulative histograms that show the total number of simulations (out of 100,000) that had locked on to the optimal arm by a certain number of steps. The network has nine agents and $p = 0.6$.

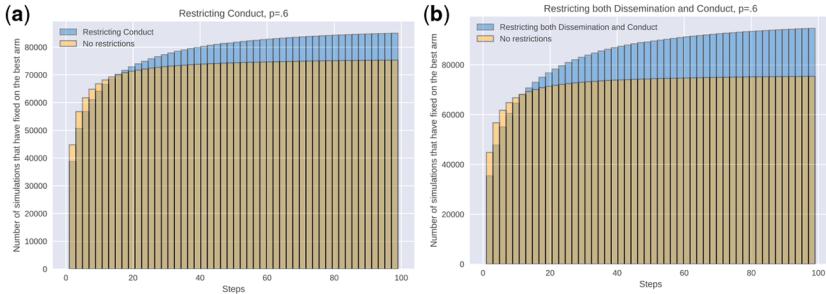


Figure 9. Cumulative histograms that show the total number of simulations (out of 100,000) that had locked on to the optimal arm by a certain number of steps. The network has nine agents and $p = 0.6$.

meaningfully slow learning. For example, Figure 10 contains a cumulative histogram showing the distributions of lock-on times when $p = 0.8$. This figure shows that restrictions on dissemination do not slow learning here. By just four steps the restricted models had overtaken the restriction-less baseline. Cumulative histograms for restrictions on just conduct as well as simultaneous restrictions on both conduct and dissemination are indistinguishable from that shown in Figure 10.

So, do restrictions slow learning? Only for some parameters. Inside the parameter range that yields the Zollman effect, successful learning is slowed, but not by much.

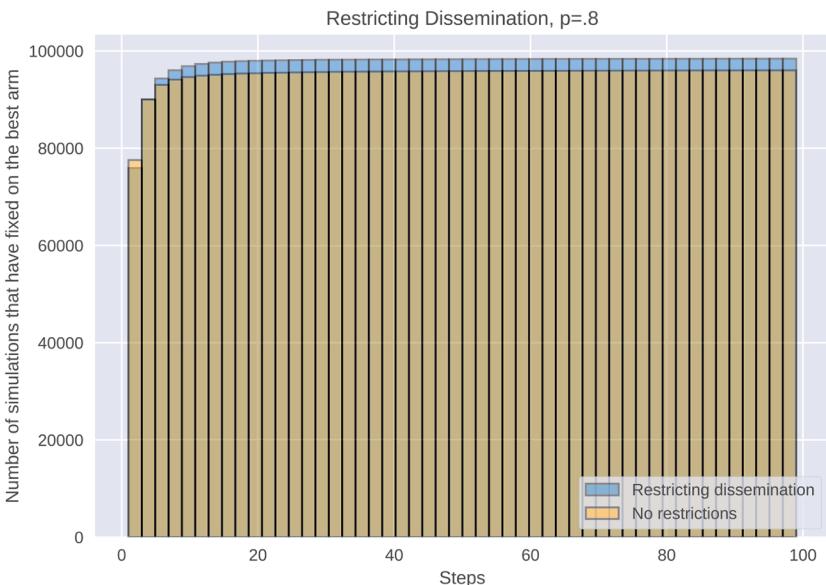


Figure 10. Cumulative histograms that show the total number of simulations (out of 100,000) that had locked on to the optimal arm by a certain number of steps. The network has nine agents and $p = 0.8$.

Restrictions do lead to a slightly longer average time-to-lock, but average time-to-lock is not too useful of a statistic because so many more simulations with restrictions successfully converge than do simulations of the restriction-less baseline. Moreover, outside of the Zollman effect range—that is, in those cases where restrictions neither increase nor decrease the network’s probability of success—restrictions do not meaningfully slow learning.

5. Conclusion

Scientists and others have a history of claiming that any restrictions on the conduct or dissemination of dual-use research will impede scientific progress. This claim is based on the widely held, and strongly felt, intuition that restrictions are counter to science. Our aim here is to probe this intuition in a systematic fashion, and our results suggest that this intuition may not be correct. Rather than impeding a scientific community’s ability to learn which of two research methodologies is superior, there is actually a range of parameters for which restrictions are conducive to successful learning. Moreover, there is no range of parameters for which restrictions on dissemination impede learning. And, as long as not too many agents have their conduct restricted, there is no range for which restrictions on conduct impede learning.

In other words, our results suggest that allowing researchers unrestricted access to dual-use methods and dual-use results does not yield any epistemic benefit. But recall that dual-use research has the potential to cause grave harm. Therefore, scientists and others who criticize proposals to restrict dual-use research ought to show not only that a lack of restrictions yields epistemic gains, but also that those gains outweigh the risks to human health. In so far as our model suggests that there are no epistemic costs to restrictions, a key component of the argument against restrictions on dual use research has been undermined.

Of course, the models described in this article are simple and highly idealized. Despite these limitations, they are not too far-fetched, especially when thought of as an idealization of the development of new vaccine candidates for pathogens governed by the select agent programme. The research groups that do this work have limited funds, and each group can spend its funds on one of two broad techniques for generating new vaccine candidates: they can either generate inactivated vaccines or they can modify live select agents. It is not known which of these methods is best for each select agent and host, so researchers experience a tension between exploring all methods and exploiting the method that, given current beliefs, appears best. In short, select agent vaccine researchers face a bandit problem. Our results suggest that restrictions on dissemination and conduct will not hinder the scientific community and may actually aid the production of successful vaccines.

Interestingly, although both restrictions on dissemination and conduct can be conducive to learning, the two types of restrictions have somewhat different effects. Restricting dissemination is more conducive to learning than restricting conduct. And

restricting conduct can hinder learning if too many agents in the scientific community are restricted. This difference is due to the fact that whereas limits on dissemination merely reduce the information that flows throughout the community, limits on conduct reduce the information that can, in principle, be gained about the bandit's arms. This somewhat nuanced understanding of the differing effects of the different forms of restrictions is not present in the extant literature on the ethics and practicality of limiting dual-use research. It is only by moving beyond intuitions to formal models that such distinctions can be recognized.

Although we motivated our models with a discussion of vaccine development for select agents, this framework fits other dual-use areas as well. For instance, some aspects of nuclear fusion research appear to resemble bandit problems. Researchers seeking more efficient ignition systems for nuclear fusion may choose between studying direct or indirect laser ignition. The latter generates information with proliferation potential, while the former does not (National Research Council [2013]). It is an open question as to whether the direct or indirect approach is more likely to succeed. Restrictions on the conduct and dissemination of ignition research can therefore be captured by the models developed in this article.

One might worry that these results recommend restricting the conduct and dissemination of science in general—not just in the case of dual-use research. After all, if learning can sometimes be improved by restricting dual-use research, why wouldn't that lesson apply more broadly to all scientific research? Here are two reasons to resist this conclusion. First, the epistemic benefits seen here might not arise outside the context of dual-use science. For instance, while we did not model the effect of 'credit' allocation (more on this below), there is reason to believe that restrictions on dissemination would disincentivize agents from pursuing the restricted approach. These credit allocation effects aren't so worrisome in the context of dual-use research, because dual-use methods are often funded by, or institutionalized within, special national security programmes that provide a steady stream of funding and jobs for scientists working on dual-use topics. But the disincentive to pursue restricted methods may be much more pronounced in the general scientific community where special streams of funding and employment do not exist. Second, even if restrictions did result in epistemic benefits, those restrictions may not be ethically justifiable in non-dual-use cases. Restrictions on the dissemination of scientific results are a form of censorship, and thus demand weighty reasons in their favour. While there is a weighty, non-epistemic reason for restrictions in the dual-use case—the clear risk of grave harm—it is not clear that the epistemic benefits of such restrictions would be sufficient to license censorship of science that does not pose a grave risk of harm.

These models could be extended in various ways. First, the incentive structure that our agents face is relatively simple. In our model each agent is presumed to be a Bayesian learner whose goal is simple: maximize the immunogenicity of the vaccines produced. Yet real scientists aren't only motivated by humanitarian aims, and the 'payouts' they receive from each experiment are not solely determined by the quality

of the vaccines they produce. Instead, scientists are also strongly motivated by a system of credit that places a high premium on publication of results (and awards priority to individuals who make large improvements over existing vaccines). Restrictions on the dissemination of certain experimental results may therefore reduce the incentive that agents have to pursue that arm (that is, scientists may be less likely to pursue a project if they know that the results will not be published). Palese ([2012], p. 115) raises this concern about research into H5N1 avian influenza, noting that ‘We need more people to study this dangerous pathogen, but who will want to enter a field in which you can’t publish your most scientifically interesting results?’. An extension of our model could explore the chilling effect of publication restrictions.

Second, our model assumes a fixed population of agents over time. In most scientific disciplines, however, the capacity to conduct research is dependent on the success of funding proposals, and these are in turn dependent on a track record of successful results. In this respect, scientists who consistently experience negative results (that is, in our toy case, produce vaccines with low immunogenicity), may not only switch methods but may be forced to exit the research network altogether. This could be modelled in our framework by replacing agents in the network if they receive a specified number of ‘failed’ payouts.

Third, the arms available to each agent are fixed and each arm yields payouts that are independent and identically distributed. This is not quite how scientific investigation works. As pointed out by a referee, different labs may have different success rates even when applying the same methods to the same pathogen. And over time, those methods may be tweaked or new methods may be invented. A more nuanced model may therefore build off Alexander’s ([2013]) variable arm bandit problems in which new arms may be discovered by an agent and disseminated as options to that agent’s neighbours.

Nonetheless, although these extensions would deepen the verisimilitude of the model, we do not expect that they would alter the central finding of this article. By providing a simple model of restrictions on dual-use research we have laid the foundation for a systematic study of the intuitive claims made by both sides in the dual-use research policy debate. This model shows that, contrary to a common view shared by scientists, policy makers, and others, restrictions on the conduct and dissemination of dual-use research do not impede the pace or reliability of that research. This suggests that restrictions on dual-use research should not be reflexively opposed due to fears that they will impede scientific progress.

Acknowledgements

The authors wish to thank audiences at the 2016 Philosophy of Science Association Meeting, and three anonymous reviewers for their comments on earlier drafts of this article.

Elliott Wagner

Department of Philosophy

Kansas State University

Manhattan, KS, USA

eowagner@ksu.edu

Jonathan Herington

Department of Philosophy

Kansas State University

Manhattan, KS, USA

jherington@ksu.edu

References

- Alexander, J. M. [2013]: ‘Preferential Attachment and the Search for Successful Theories’, *Philosophy of Science*, **80**, pp. 769–82.
- Bala, V. and Goyal, S. [1998]: ‘Learning from Neighbours’, *The Review of Economic Studies*, **65**, pp. 595–621.
- Baltimore, D., Berg, P., Botchan, M., Carroll, D., Charo, R. A., Church, G., Corn, J. E., Daley, G. Q., Doudna, J. A., Fenner, M., Greely, H. T., Jinek, M., Martin, G. S., Penhoet, E., Puck, J., Sternberg, S. H., Weissman, J. S. and Yamamoto, K. R. [2015]: ‘A Prudent Path Forward for Genomic Engineering and Germline Gene Modification’, *Science*, **348**, pp. 36–8.
- Boulanin, V. and Verbruggen, M. [2017]: ‘Mapping the Development of Autonomy in Weapon Systems’, Technical Report, Stockholm International Peace Research Institute.
- Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, G. C., Steinhardt, J., Flynn, C., higeartaigh, S., Beard, S., Belfield, H., Farquhar, S., Lyle, C., Crootof, R., Evans, O., Page, M., Bryson, J., Yampolskiy, R. and Amodei, D. [2018]: ‘The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation’, Technical Report, available at <arxiv.org/pdf/1802.07228.pdf>.
- Buchanan, A. and Kelley, M. C. [2013]: ‘Biodefence and the Production of Knowledge: Rethinking the Problem’, *Journal of Medical Ethics*, **39**, pp. 195–204.
- Bush, V. [1945]: *Science, the Endless Frontier: A Report to the President, July 1945*, Washington, DC: US Government Printing Office, available at <www.nsf.gov/od/lpa/nsf50/vbush1945.htm>.
- Casadevall, A., Howard, D. and Imperiale, M. J. [2014]: ‘An Epistemological Perspective on the Value of Gain-of-Function Experiments Involving Pathogens with Pandemic Potential’, *Mbio*, **5**, pp. e01875–14.
- Chambers, M. A., Graham, S. P. and Ragione, R. M. L. [2016]: ‘Challenges in Veterinary Vaccine Development and Immunization’, in T. Sunil (ed.), *Vaccine Design: Vaccines for Veterinary Diseases*, New York: Humana Press, pp. 3–35.
- Douglas, T. [2014]: ‘The Dual-Use Problem, Scientific Isolationism, and the Division of Moral Labour’, *Monash Bioethics Review*, **32**, pp. 86–105.
- Evans, N. G. [2013]: ‘Great Expectations: Ethics, Avian Flu, and the Value of Progress’, *Journal of Medical Ethics*, **39**, pp. 209–13.

- Faburay, B., LaBeaud, A. D., McVey, D. S., Wilson, W. C. and Richt, J. A. [2017]: 'Current Status of Rift Valley Fever Vaccine Development', *Vaccines*, **5**, p. 29.
- Fouchier, R. A. M. [2015]: 'Studies on Influenza Virus Transmission between Ferrets: The Public Health Risks Revisited', *Mbio*, **6**, pp. e02560–14.
- Franceschini, G., Englert, M. and Liebert, W. [2013]: 'Nuclear Fusion Power for Weapons Purposes', *The Nonproliferation Review*, **20**, pp. 525–44.
- Grim, P. [2009]: 'Threshold Phenomena in Epistemic Networks', in M. Hadzikadic and T. Carmichael (eds), *Complex Adaptive Systems and the Threshold Effect: Views from the Natural and Social Sciences*, Menlo Park, CA: AAAI Press, pp. 53–60.
- Herfst, S., Schrauwen, E. J. A., Linster, M., Chutinimitkul, S., de Wit, E., Munster, V. J., Sorrell, E. M., Bestebroer, T. M., Burke, D. F., Smith, D. J., Rimmelzwaan, G. F., Osterhaus, A. D. M. E. and Fouchier, R. A. M. [2012]: 'Airborne Transmission of Influenza a/H5N1 Virus between Ferrets', *Science*, **336**, pp. 1534–41.
- Holman, B. and Bruner, J. P. [2015]: 'The Problem of Intransigently Biased Agents', *Philosophy of Science*, **82**, pp. 956–68.
- Imperiale, M. J. and Casadevall, A. [2014]: 'Vagueness and Costs of the Pause on Gain-of-Function (GOF) Experiments on Pathogens with Pandemic Potential, Including Influenza Virus', *Mbio*, **5**, pp. e02292–14.
- Kaiser, D. and Moreno, J. [2012]: 'Dual-Use Research: Self-Censorship Is Not Enough', *Nature*, **492**, pp. 345–7.
- Kummerfeld, E. and Zollman, K. J. S. [2016]: 'Conservatism and the Scientific State of Nature', *British Journal for the Philosophy of Science*, **67**, pp. 1057–76.
- Lentzos, F. [2008]: 'Countering Misuse of Life Sciences through Regulatory Multiplicity', *Science and Public Policy*, **35**, pp. 55–64.
- Lipsitch, M. and Inglesby, T. V. [2014]: 'Moratorium on Research Intended to Create Novel Potential Pandemic Pathogens', *Mbio*, **5**, pp. e02366–14.
- Manini, I., Trombetta, C. M., Lazzeri, G., Pozzi, T., Rossi, S. and Montomoli, E. [2017]: 'Egg-Independent Influenza Vaccines and Vaccine Candidates', *Vaccines*, **5**, p. 18.
- Mayo-Wilson, C., Zollman, K. and Danks, D. [2013]: 'Wisdom of Crowds versus Group-think: Learning in Groups and in Isolation', *International Journal of Game Theory*, **42**, pp. 695–723.
- Meeusen, E. N. T., Walker, J., Peters, A., Pastoret, P. -P. and Jungersen, G. [2007]: 'Current Status of Veterinary Vaccines', *Clinical Microbiology Reviews*, **20**, pp. 489–510.
- Miller, S. and Selgelid, M. J. [2008]: *Ethical and Philosophical Consideration of the Dual-Use Dilemma in the Biological Sciences*, Amsterdam: Springer.
- National Institutes of Health. [2014]: 'Tools for the Identification, Assessment, Management, and Responsible Communication of Dual Use Research of Concern', Technical Report, United States Government, available at <www.phe.gov/s3/dualuse/Documents/durc-companion-guide.pdf>.
- National Research Council. [2013]: *Assessment of Inertial Confinement Fusion Targets*, Washington, DC: National Academies Press.
- National Science Advisory Board for Biosecurity. [2012]: 'Full Recommendations of NSABB Regarding March 29–30, 2012 Meeting to Review Revised Manuscripts on Transmissibility of A/H5N1', Technical Report, available at <www.nih.gov/sites/default/files/about-nih/nih-director/statements/collins/03302012_NSABB_Recommendations.pdf>.

- Palese, P. [2012]: ‘Don’t Censor Life-Saving Science’, *Nature*, **481**, p. 115.
- Palese, P. and Wang, T. T. [2012]: ‘H5N1 Influenza Viruses: Facts, Not Fear’, *Proceedings of the National Academy of Sciences*, **109**, pp. 2211–3.
- Rosenstock, S., Bruner, J. and O’Connor, C. [2017]: ‘In Epistemic Networks, Is Less Really More?’, *Philosophy of Science*, **84**, pp. 234–52.
- Selgelid, M. J. [2016]: ‘Gain-of-Function Research: Ethical Analysis’, *Science and Engineering Ethics*, **22**, pp. 923–64.
- Villar, S. S., Bowden, J. and Wason, J. [2015]: ‘Multi-armed Bandit Models for the Optimal Design of Clinical Trials: Benefits and Challenges’, *Statistical Science*, **30**, pp. 199–215.
- Zollman, K. J. S. [2007]: ‘The Communication Structure of Epistemic Communities’, *Philosophy of Science*, **74**, pp. 574–87.
- Zollman, K. J. S. [2010]: ‘The Epistemic Benefit of Transient Diversity’, *Erkenntnis*, **72**, pp. 17–35.

