How Should We Promote Transient Diversity in Science?

Jingyi Wu and Cailin O'Connor

August 10, 2021

Abstract

Diversity of practice is widely recognized as crucial to scientific progress. If all scientists perform the same tests in their research, they might miss important insights that other tests would yield. If all scientists adhere to the same theories, they might fail to explore other options which, in turn, might be superior. But the mechanisms that lead to this sort of diversity can also generate epistemic harms when scientific communities fail to reach swift consensus on successful theories. In this paper, we draw on extant literature using network models to investigate diversity in science. We evaluate different mechanisms from the modeling literature that can promote transient diversity of practice, keeping in mind ethical and practical constraints posed by real epistemic communities. We ask: what are the best ways to promote the right amount of diversity of practice in such communities?

1 Introduction

In northern Australia, locals have strategically incorporated controlled burning in their farming practices. They do so for various reasons—to favor certain types of vegetation to attract prey, to stimulate the growth of crops, or to restore biodiversity (Jones, 2012; Bonta et al., 2017). A central variable in the fire management practices—that certain species of fire-foraging raptors can intentionally carry burning sticks to spread fire in order to flush out prey—was widely known to local Indigenous populations through a combination of direct observation and retelling of stories by elders (Bonta et al., 2017). But the behaviors of these raptors remain relatively unknown to, or discounted by, outsiders. Though many Indigenous rangers take into account the risks posed by raptors that could cause controlled burns to jump to new locations (also known as fire-jumping), others, including local officials, remain largely skeptical of this piece of Indigenous Ecological Knowledge (IEK), characterizing the raptors' behaviors as accidental or unintentional (Bonta et al., 2017). Local officials often blame firefighters' inexperience when a fire jumps to new locations, even when firefighters themselves attribute the fire-jumping to the birds (Bonta et al., 2017, 713). Many have argued, for this reason, that incorporating and accepting IEK within a larger scientific and regional planning community is essential to fire management in the area (Russell-Smith et al., 2013; Bonta et al., 2017).

It has been widely argued that a diversity of practices is crucial to scientific progress. If all scientists perform the same tests, they might miss important insights that other tests would yield. If all scientists adhere to the same theories, they might fail to explore other options which, in turn, might be superior. For this reason, many have pointed out that it is worth promoting and preserving a diversity of beliefs and practices within scientific communities. In the case just discussed, IEK adds to the diversity of beliefs and practices in fire management communities in Australia. This diversity arguably improves the success of such communities, by introducing theories—such as those related to avian fire spreading—that may ultimately prove very useful in fire management.

Our goal in this paper is to discuss proposals for how to go about promoting beneficial diversity of scientific practice, drawing on extant literature. In particular, we focus on one subset of literature in philosophy of science—that using network models to investigate the benefits of diversity in science, and exploring mechanisms that promote such diversity. There is a wide ranging qualitative literature on this topic with a deep history. Our contribution here focuses only on this smaller body of work using models to think about the problem. Our aim is to see what suggestions and proposals can be drawn from this literature, and how they can inform our thinking about promoting diversity of practice. Of course, this overview will be just one piece of the puzzle in thinking about why diversity is beneficial, and considering the best ways to achieve it.

As will become clear, in this paper we focus on diversity related to the *practices* of science. This sort of diversity is present when scientists vary their activities in ways that allow for a broader exploration of scientific possibilities. We will focus on models where actors have the options to favor different theories, and thus to try different tests. These models consider how/when it benefits them to do so, and what can lead them to wider or narrower exploration. Because belief and action are tightly associated, we are also interested in diversity of the sorts of beliefs and assumptions that, in turn, generate various practices.[1] It has been widely argued that increasing demographic/personal diversity is one way to increase diversity of beliefs in science since those with different backgrounds will tend to bring different beliefs, assumptions, and interests to their practice (Haraway, 1989; Longino, 1990; Fehr, 2011). The models we discuss primarily focus on other factors promoting diversity of practice. This said, as will become clear, there are connections between the modeling work we discuss and work on the importance of demographic diversity in science.

The paper will proceed as follows. In section 2 we discuss in more detail why diversity of practice is crucial to scientific progress. In particular we discuss the idea of "transient diversity"—that a successful scientific community will have a period of sufficient exploration to test many plausible theories and options. In this section we present and overview a number of modeling results

[1] This is sometimes called cognitive diversity or epistemic diversity in science (Fehr, 2011).

showing why transient diversity is important, and how to generate it. Section 3 considers a less popular topic of exploration—how diversity of practice can be harmful. Ultimately some scientific practices are better than others. While it is important to explore many possibilities, it can also be inefficient to spend too much effort on sub-optimal theories and practices. In section 4 we present the main contribution of the paper: an in-depth discussion of concrete proposals for maintaining the right levels of diversity in science, keeping in mind practical and ethical constrains. We do three main things in this section. 1) We consider how scientific communities might mimic decision theoretic norms for exploration. I.e., after identifying optimal levels of diversity and exploration in science we ask: is it possible to replicate these optimal levels in real communities? 2) We consider how we might instead incrementally improve existing scientific communities. I.e., we ask: starting from the status quo, what practical changes might be implemented for us to ensure and gain from transient diversity in science? And 3) we consider how mechanisms from the modeling literature for promoting transient diversity might work in real communities. Which mechanisms are most promising? Which are not likely to benefit epistemic groups? Section 5 concludes.

2 The Benefits of Transient Diversity in Science

In order for a scientific community to settle on successful and pragmatically useful theories, the community typically must first explore some diversity of possibilities. If not, the group may fail to ever seriously consider highly successful theories and practices, and instead preemptively settle on some relatively poor alternative. This is sometimes referred to as a period of transient diversity in science (Zollman, 2010).

This point has been made many times in the philosophy of science. Kuhn (1977) praises disagreement in science, pointing out that it is necessary to encourage exploration of rival theories. For this reason he argues that a diversity of inductive standards is permissible. Both Kitcher (1990) and Strevens (2003) recognize the importance of division of labor, or a focus on different problems by different community members, in science. As they point out, if all scientists work on the same problems important insights might be missed.[2] Smaldino and O'Connor (2021) point out that disciplinary structure in science can help protect a diversity of methods in the face of human tendencies towards conformity. Although they argue that interdisciplinary contact is important to the spread of good scientific methodology, they advocate for the protection of disciplinary structure for this reason.

Recently in philosophy of science, a number of authors have used the "network epistemology" paradigm to explore 1) the benefits of transient diversity in scientific communities and 2) how such benefits might be achieved. In the

[2]They use models to explore how this division of labor can be promoted even when scientists tend to agree on which problems are most promising. Both authors argue that credit incentives might promote a diversity of approaches in science, and thus might be a good thing.

rest of this section we will describe results from this (and related) frameworks, which will provide a concrete starting point for further discussion.

Zollman (2007, 2010) uses network models to explore the emergence of scientific consensus. The particular models he employs are drawn from the work of Venkatesh and Goyal (1998) in economics. The models assume a network of scientists where edges represent communicative ties. Scientists in the model face a problem of selecting between several different action-guiding theories, one of which will be better than the rest. In particular, the actors attempt to solve "multi-armed bandit" problems, so named for their similarity to slot machines (or "bandits"). Individuals may choose between different options (or arms), i. These options have different characteristic probabilities of success, p_i. The goal is to choose the most successful option. But there is a trade-off between exploring—taking time to examine each arm carefully to learn its rate of success—and exploiting—actually taking the most successful action and reaping the benefits of doing do. The arms here can represent practices in science that yield epistemic successes at different rates. For instance, the practice might involve trying two different drug treatments on patients with liver disease, or believing Lyme disease can or cannot exist in a chronic state and acting accordingly (O'Connor and Weatherall, 2018), or acting to either explore or ignore the dangers of smoking (Weatherall et al., 2020).

In the models presented by Zollman, actors have credences about which option is best. They use these credences to select between arms, and upon doing so observe the success or failure of their choice. Actors then use these observations to update their credences, so that over time they will generally come to learn more about the success rates of the arms. Furthermore, agents also update their beliefs based on the actions and observations of their network neighbors. In this way, actors can learn about actions that they themselves did not take. Typical versions of the model assume that agents update using Bayes rule, and thus adhere to standards of rationality when using data to change beliefs. This process might reflect one, for example, where some scientists come to suspect that tobacco smoking is dangerous and alter their practices to test this theory. Over time their results might make them, and also their scientific peers, more confident that tobacco is harmful, thus leading to further exploration of this possibility.

In these models the social influence inherent in the network structure means that groups tend towards consensus. Typically enough actors gather data about the best option that the entire group eventually ends up accurately believing that it is, indeed, best. But in many versions of these models communities can also fail to form a good consensus. Zollman assumes that actors in the models myopically choose whichever option they currently believe is most successful. This might correspond to a gambler playing the bandit arm she likes best, or a scientist generally testing the theory they find most promising, or a doctor prescribing only the medication she thinks most efficacious. In some cases, a string of misleading data can lead an entire community to prefer a suboptimal option. Once the entire group focuses on this option, they stop testing other ones and settle on a poor consensus. Zollman (2010) gives a case study exemplifying this latter possibility. In the early 20th century, scientists debated whether stomach acid or bacteria was the primary cause of peptic ulcer disease. A highly influential study by Palmer (1954) convinced the research community that bacteria could not live in the stomach, resulting in a consensus on the acid theory. This research was flawed, but was only finally overturned by the work of Warren and Marshall (1983).

We can now make clear how these models connect with other thinking on the importance of diversity of practices in science. A community in these sorts of models can fail if it does not spend enough time testing all the possible arms. And given that beliefs about the arms shape which actions scientists actually try, diversity of belief is key to ensuring diversity of practice.

Zollman (2007, 2010) focuses on the role of communication structure in preserving transient diversity of beliefs. As he shows, less connected networks, i.e., those where fewer individuals communicate, are more likely to end up at the correct consensus. In these less connected groups it is more likely that pockets of diverse beliefs are preserved long enough for a good consensus to emerge. In more tightly networked groups, misleading data is more likely to sway the entire community to settle on a sub-optimal theory. The counter-intuitive suggestion is that because transient diversity of beliefs is so important, there may be situations where it is better for scientists to communicate poorly simply to preserve this diversity. For example, if the work of Palmer (1954) had been less influential, researchers unfamiliar with it might have continued to explore bacterial causes for ulcer disease, potentially leading to quicker confirmation of the more accurate theory.

Several other lines of investigations have found deeply similar results. One relevant body of literature, focusing on cultural innovation and problem solving, explores models where actors try to solve NK landscape problems. These problems involve searching a solution space with multiple "peaks" so that sometimes actors get stuck at local optima despite the presence of better global solutions.[3] March (1991) first identified NK landscapes as a good way to capture group innovation, and showed that too much fast social learning could lead groups to converge to local optima and fail to discover better solutions. This is analogous to a case where a tightly connected scientific community fails to explore many theories, and preemptively settles on one. Lazer and Friedman (2007) and Fang et al. (2010) find that agents in less connected networks tend to find better solutions to these problems (though it takes them longer). Derex and Boyd (2016) present a cleverly designed experiment meant to test these results where teams of six solve problems wherein they can accumulate technological advances by building on earlier stages. The learning set-up is one where a failure to engage in enough early exploration eliminates possibilities later on. They find that highly connected groups, where all six members see each innovation in their group, fail to discover the highest level technologies. Less connected groups are much more likely to do so, as small teams explore a diversity of paths through the problem,

[3]This area of research connects back to early work in biology by Wright (1932) (see a discussion in Fang et al. (2010)). As Wright pointed out, biological populations benefit from a partially isolated subgroup structure that preserves a diversity of adaptations.

and combine their discoveries.[4]

Besides social disconnection, extensions to the bandit model explore a number of other mechanisms that can also ensure a diversity of practice, and thus improve community success. Zollman (2010) points out that intransigence or stubbornness on the part of individual scientists can likewise preserve such diversity. If scientists are unwilling to revise their beliefs about the success of various theories, this can lead individuals to keep testing a seemingly unpromising option long enough to discover its true merits. This suggestion relates to claims by Kuhn that different inductive standards are acceptable for science. If some scientists are "irrationally" stubborn, this might still benefit the community since they insist on testing a diversity of possibilities rather than following only the most promising ones.

Relatedly Gabriel and O'Connor (2021) consider models where individuals have a tendency towards confirmation bias, i.e., where they are more likely to engage with evidence that fits their prior beliefs. Like stubbornness, this feature leads scientists to stick with theories longer than they would if they were behaving in a strictly rational manner, which preserves group diversity of beliefs and benefits the community. This suggests that perhaps a seemingly harmful reasoning bias actually is beneficial to social learning where a diversity of beliefs and practices can be helpful.[5]

Kummerfeld and Zollman (2015) show how individual tendencies towards exploration, where scientists continue to test theories that they personally suspect are suboptimal, will likewise preserve diversity of practice and improve community performance. As they argue, though, scientists are not generally incentivized to explore at a socially optimal level, yielding a free rider problem. They take this as a reason that institutions such as funding bodies and award giving agencies might play a beneficial role in promoting exploratory science.

Wu (2021b) finds a similar effect through yet another mechanism. She considers situations where some dominant individuals in an epistemic community systematically undervalue or ignore the testimony of some marginalized individuals. This might represent a case where members of one racial or ethnic group ignore those in another, or members of one scientific discipline tend to devalue contributions from another. This addition to the model is inspired by Fricker (2007)'s concept of "testimonial injustice" and Dotson (2011)'s concept of "epistemic quieting." As Wu shows, this can lead to a surprising epistemic advantage for those in the marginalized community, which tends to reach accurate beliefs more often than communities without testimonial injustice. This is because they update their beliefs on data from the entire network. Furthermore, the dominant group receives less data overall, meaning they tend to spend more time testing a wider variety of, possibly unpromising, theories. The marginalized group can learn from this diversity of practice, while those ignoring their

[4]See also Mason et al. (2008); Jönsson et al. (2015) who yield similar findings. Mason and Watts (2012), on the other hand, find a general disadvantage to low connectivity in problem solving groups.

[5]This finding fits with informal claims from Mercier and Sperber (2017) that confirmation bias might play a functional role in groups.

out-group cannot.[6]

There is one further thing to mention about all of these models, which is that generally factors which improve consensus on better theories in these models also slow down the emergence of community consensus. This should make intuitive sense. If scientists are exploring more theories long enough to really get a sense of their merits, they will not be swiftly settling on one theory.

Altogether this body of literature strongly supports the intuitive claim that it is indeed important to promote a transient diversity of beliefs and practices in scientific communities. As we have seen, though, there are many factors that might lead to this sort of diversity: reduced network connectivity, individual stubbornness (or different standards for induction), confirmation bias, the active promotion of exploratory or risky science by funding bodies, and even testimonial injustice.

In the next section, we will turn to discuss another side of the picture here—the ways that transient diversity can be harmful to a scientific community. From there we will move on to discuss how this literature should inform our thinking about real communities.

3 The Harms of Transient Diversity in Science

The work described in the last section indicates that a transient diversity of approaches is crucial to ensuring that scientific communities do not miss out on promising theories. But there is a tension inherent in the promotion of this sort of diversity. There is a cost to using suboptimal theories, paradigms, and methods in the sciences. They are suboptimal, and thus do not embody the current best approaches to action and investigation. This is especially clear when it comes to areas like medicine, where incorrect beliefs can have direct negative impacts on patients. For instance, doctors who continue to explore the theory that "cigarettes promote health" will have direct negative health impacts. But even in other areas adherence to a poor theory can impede progress and create inefficiencies in science. The tension here is exacerbated by the observation that in science there are typically limited resources. Researchers do not have the time, money, or energy to explore multiple options indefinitely.

These observations are related to the explore/exploit trade-off inherent to bandit problems. In order to find out about a bandit arm, you have to pull it, and you have to pull it enough that you get a decent sample of outcomes. In order to learn about new possibilities in science, you likewise have to test them.[7] This means that to promote good learning, there must be periods of inefficiency or poor outcomes. This raises a question for scientific communities, though: how can a group maximize the benefits of transient diversity while minimizing

[6] Wu (2021b) points out that her findings support a version of standpoint epistemology, where a disadvantaged social status can lead to epistemic advantages.

[7] Sometimes old tests are able to illuminate new theories as well as old ones. In the conclusion we will say a bit more about the sorts of problems in science that are not well modeled by bandit problems for this reason.

the harms? What is the best way to ensure that promising theories are duly tested, while avoiding the costs of using suboptimal options for too long?

Decision theory provides normative solutions when it comes to individuals engaged in bandit problems (Gittins, 1979; Lai and Robbins, 1985; Gittins et al., 1989; Berry and Fristedt, 1985). These solutions identify the optimal amount of exploration to ensure actors eventually settle on the best arm, but do not waste too much effort exploring. Although these solutions are sometimes complex, approximate solutions exist that are not hard to implement. There are a variety of successful *greedy* strategies, for instance, that typically select the best option based on past observation but with some small probability explore other options (Sutton and Barto, 2018). Some of these strategies decrease the level of exploration over time so that it starts hotter and slows down. It has been shown that by employing these greedy strategies one can dependably learn to pick the best option, while also spending most of the time implementing successful strategies.

We might suggest that science attempt to mimic these normative solutions. When it comes to scientific communities, though, the issue is that they are not run by single decision makers or even by centralized bodies. Instead individual scientists (or collaborative groups of scientists) make decisions about what to work on, when, and for how long, in a relatively uncoordinated way. This, of course, is part of why we saw a number of suggestions in section 2 that appealed to properties of individuals and communities that might promote transient diversity of approach without a coordinated decision maker: stubbornness, credit motivation, low connectivity, confirmation bias, etc.

But while these sorts of solutions to the transient diversity problem can indeed improve final outcomes, many of them are quite inefficient when compared to ideal approaches to bandit problems. Rosenstock et al. (2017), for instance, further consider the sort of models presented by Zollman (2007, 2010), and show that the benefits of decreased communication hold only in cases where the problem is especially difficult.[8] (When the problem is easy, on the other hand, almost every community successfully solves it, meaning there is no benefit to limiting communication.) As they point out, this makes policy suggestions from Zollman (2007, 2010) sound almost paradoxical. The proposal is to decrease the flow of information in the community in exactly those situations where good data is hard to gather. While this sort of decreased information flow does improve eventual outcomes, it does so at a huge cost to efficiency—diversity is only preserved by preventing actors from learning from useful data for a significant length of time. As Rosenstock et al. (2017) argue, we should try to avoid this massive inefficiency, while still exploring options for a sufficient length of time.

There is another potential downside to mechanisms that promote transient diversity, which is that across models they also tend to lead to polarization. When these mechanisms slow learning too far they can lead to situations where disagreement in the community becomes stable, rather than transient. This

[8]This reflects findings on NK landscape models as well. In very simple landscapes, like those with single peaks, decreased communication does not improve group success. Derex and Boyd (2016) also find empirically that limited group communication only improves performance when groups are faced with complex problems.

stability of disagreement prevents the group from ever converging to a good outcome, and leaves some individuals continuing to test poor theories. Zollman (2010), for example, points out that when individuals are stubborn *and* when group connectivity is low in his models learning becomes so slow that the community mimics one that is failing to come to consensus. Gabriel and O'Connor (2021) point out that while moderate levels of confirmation bias increase the chances of group success, higher levels of confirmation bias instead lead to polarization. Sub-groups form where individuals in each one only listen to the sets of evidence that fit with their current beliefs. This might track a case where, for example, one group is convinced that hydroxychloroquine is a successful COVID treatment, and the other that it is not. Each group only considers research supporting their position, thus failing to end a period of transient diversity of belief. In the models from Wu (2021b) looking at testimonial injustice, there are cases where marginalized groups reach accurate beliefs, but dominant groups who ignore them continue to prefer an inaccurate theory. In such cases, there are potential harms that could arise from the dominant group failing to ever adopt more successful practices. Consider the example about fire-spreading raptors that we started this paper with. Here, outsiders and local officials' resistance to Indigenous Ecological Knowledge arguably leads to miscalculations regarding whether a fire is under control, thus endangering the local region with potential catastrophes. These models all track cases where too much diversity of practice, lasting for too long, is a bad thing.

In special situations, there are further risks to diversity of practice that arise from industrial and political propagandists attempting to influence scientific beliefs. Such propagandists can take advantage of doubt, uncertainty, and lack of consensus in scientific communities to delay action against public health risks like tobacco smoking and the use of fossil fuels (Oreskes and Conway, 2011; O'Connor and Weatherall, 2019; Weatherall et al., 2020). For instance, tobacco interests funded research on asbestos in order to create doubt about whether tobacco smoke was a main cause of lung disease. Big tobacco also widely shared legitimate, independent research arguing that smoking might be safe, again to fuel such doubt (Oreskes and Conway, 2011). Holman and Bruner (2017) point out that when scientists employ a diversity of methods, industry can fund just those scientists whose findings tend to support industry interests and thus shape epistemic progress. In all these cases, diversity of beliefs and practice in science are weaponized to create illegitimate dissent. We can think of these propagandists as extending the period during which less-successful practices continue to be used beyond what is necessary to develop good beliefs. In doing so, they increase the harms that result from transient diversity of practice.

4 Moving Forward

Given what we have seen so far, a question arises: just what are the best ways to promote cognitive diversity in science, given our understandings of the harms, benefits, and mechanisms of promoting transient diversity? The goal of the rest of this paper will be to shed some light on this question. In doing so, we need to keep in mind a few sorts of considerations. 1) We want to promote enough diversity to ensure good outcomes, while minimizing harms from testing suboptimal theories. 2) We recognize that there are many facts about scientific communities that constrain the sorts of solutions to this problem that might be effective. And relatedly, 3) there are ethical considerations at play in thinking about the best ways to ensure transient diversity of beliefs.

While there are many ways to structure the discussion that follows, we divide it into three main parts. We start with optimal solutions identified in the decision theory literature and ask: in what ways can real scientific communities mimic these solutions? (Or not?) We then move on to ask: starting from real scientific communities, what smaller changes might be made to promote better levels of transient diversity? And last, we assess the usefulness of various mechanisms to promote transient diversity introduced in section 2. We wrap up the section by turning back to NK-landscape problems and discussing cases where more dramatic diversity of practice might be beneficial.

4.1 Approximating Ideal Solutions

Let us start with the normative recommendations of decision theory for individuals facing bandit problems. In particular, this means we will further consider the set of greedy strategies described in the last section. In these strategies, most of the time an individual focuses on the most promising possibility, and then tests others with some small probability (either fixed or decreasing over time). If a scientific community could be shaped where individuals were perfectly able to coordinate behavior, and perfectly able to communicate results with each other, perhaps that community would be able to mimic these ideal strategies. They could decide on a plan to divide the labor of investigation either such that each scientist would test alternatives with a small probability, or such that a small group of labs would always test less promising alternatives and communicate their findings to the larger group. Such a community would be guaranteed to ensure enough transient diversity of practice, while still exploiting successful practices. Notice that under this proposal cognitive diversity is not required to promote transient diversity of practice. Instead, the necessary diversity of practice is ensured by community agreement to do so.

Of course, as noted in the last section, this is not practical for most real epistemic communities for a number of reasons. Most pressing is the fact that individual scientists make their own decisions about what topics to investigate. These decisions are driven by a wide set of factors including prior beliefs about which theories are promising (Kuhn, 1977; Zollman, 2007), credit incentives (Kitcher, 1990; Strevens, 2003), funding constraints, what topics are popular among colleagues and members of the public, etc. Central coordinating bodies cannot simply hand topics out to scientists and demand that they investigate them. In this sort of regime coordination is difficult to achieve.

One approximation might be promoted by funding bodies, which, by selecting projects to fund, rather than doling out topics, might be able to shape the overall exploratory tendencies of a community. In such a regime, maybe most money is devoted to the most promising theories, but smaller pots are systematically devoted to less promising options. Relatedly, some philosophers have recently grappled with the importance of high-risk/high-reward science, and have argued that funding bodies should increase funding to the the sort of science that looks relatively unpromising at the moment, but may still yield important discoveries (Stanford, 2019; Currie, 2019).[9] Of course, this points to another practical constraint—which is that different scientific communities work differently, meaning that attempts to ensure the right levels of transient diversity may not work the same way in each case. In particular, some disciplines are heavily funded by one, or just a few, centralized funding sources, in which case those sources might dole out resources in an approximately efficient way. In other disciplines, there is little outside funding for projects, or else funding comes from a diverse set of sources and this sort of coordination will not be possible.

Another factor of real communities that stands in the way of matching optimal decision theoretic strategies has to do with constraints created by lab structures and other aspects of science. Nersessian (2019) discusses the ways that physical objects, as well as models and theories, constrain the practice of science. As she points out, scientific labs innovate and change, but this innovation and change is deeply shaped by the physical objects making up a lab and the conceptual resources available to it. The next step of research for some group is almost always constrained by current research projects. The "optimal" model we proposed will not be practical inasmuch as labs cannot just change gears to research entirely new topics, even if the funding is available to do so.

Again, there are differences between scientific communities here, though, since some areas have relatively small start-up costs for switching research focus, and others relatively large. For instance, in an area of research that requires expensive, specialized equipment, it may be highly difficult for labs to switch projects, while for computational modelers it might be relatively easy. Mitigating this issue is the fact that in a scientific community exploratory strategies can be distributed across a group. I.e., unlike in the decision theoretic heuristics described in the last section each scientist need not continually switch from promising to less promising research topics, as long as some small number of scientists are strictly devoted to less promising topics. In research communities where projects are highly constrained by resources, centralized funding bodies might want to be especially attentive to protecting labs that continue to explore less promising theories, and which might as a result tend to lose funding. If these labs are protected with long-term grants, or special funding measures, diversity of practice can be maintained without requiring scientists to switch topics.

Another difficulty arises vis a vis communication. In the decision theoretic

[9]Related to this are worries that there are forces in science that are inherently conservative, i.e., that push researchers away from unpromising, unusual, or high risk research projects (Kummerfeld and Zollman, 2015; Stanford, 2019; O'Connor, 2019). If so, then perhaps funding bodies might need to work extra hard to promote this sort of research.

heuristic, there is no need to communicate results, because one individual sees them all and is able to develop accurate beliefs as a result. In real communities, communication is often imperfect. While some labs may manage to communicate their results widely, others might be unable to do so, because their research is not of widespread interest, because of biases towards high prestige institutions, because of differences in the communication skills of researchers, etc. In general, it is a property of human social networks that information spreads and diffuses at different rates and to different recipients depending on its content, and who shares it (Vosoughi et al., 2018). Furthermore, many previous investigations reveal that relevant information often fails to spread in scientific communities. For instance, retractions are often cited approvingly long after they have been retracted, even though many in the community are aware that this is no longer appropriate (Neale et al., 2010; Cor and Sood, 2018). When these failures of communication happen, it may be difficult for scientific communities to approximate decision theoretic solutions because researchers may be unaware of which theories and options are, in fact, the most promising ones at any particular time. Again, this is a situation where central coordinating agencies, like grant giving bodies, may play a key role. As long as someone is aware of all the diverse sorts of research going on, and is able to track and synthesize this information, then it might be possible to coordinate exploration across the community.

There is an issue, though, that goes beyond simple constraints. There is often deep disagreement between individuals about what scientific theories are the most promising ones. Indeed, as noted in the last section, many proposals for promoting diversity of practice proceed by promoting diversity of belief. But when individuals disagree about the promise of different theories, how is exploration of the underlying space to be efficiently divided? Bandit models assume that the process of exploration is a relatively straightforward one—each success and failure is easily observed and straightforwardly comparable to past successes and failures. Scientific evidence is often not like this. There is room for substantive debate about what different evidence tells one about the world, what theories are supported by this evidence, and what sets of data are comparable. The point here is that even if there were a decision maker who could efficiently allocate labor across a community, it is often impossible to know how this decision maker ought to allocate labor given the complexities of real scientific evidence. One way to work around this problem might involve using lotteries to make funding decisions. Lotteries can ensure that a diversity of projects are funded, without requiring central coordinating groups to make hard decisions about which are most promising. We return to this idea in the conclusion.

We also need to consider the ethical costs inherent in some proposals to approximate the ideal decision theoretic strategies. As noted above, in some disciplines, practical constraints might make it difficult for individual scientists to periodically test suboptimal strategies. As a result, it might make more sense for a small subgroup of scientists to always test suboptimal strategies. However, this proposal might be unfair in so far as this small subgroup of scientists on average would receive less credit. To promote the benefits of cognitive diversity fairly, then, it might be warranted to introduce external mechanisms to compensate exploratory scientists. This relates to proposals from Stanford (2019) and Currie (2019) about using funding and credit incentives to promote this sort of research in the first place.[10]

There is another ethical worry, which relates to cases where tests themselves cause harms. This could include environmental harms—testing nuclear bombs, for instance, requires contaminating some area with radioactive material. One might argue on the basis of such harms that diversity of practice is not always worth promoting even when it might lead to important epistemic progress.

This issue arises with human subjects research such as clinical trials in medicine. Testing unpromising therapies on humans violates a widely recognized principle in research ethics called equipoise. Equipoise requires physicians to enroll patients in a clinical trial only when they are uncertain about, or equally poised between, the relative therapeutic merits of the treatments involved in the trials (Fried, 1974; London, 2009). Otherwise, physicians are ethically required to administer the better drug to the patient. This seems to be in direct conflict with the epistemic mandate to promote diversity of exploration.

However, in actual practice equipoise comes in many different formulations. One of its prominent formulations, clinical equipoise, interprets the uncertainty as a lack of consensus within the scientific community about the relative merits between treatments (Freedman, 2017; London, 2009). In a case where there is debate about these merits, scientists that test less promising alternatives would not be in violation of clinical equipoise. Moreover, even if the relevant formulation of equipoise is a personal one, i.e. individual physicians enroll patients in clinical trials if they are *personally* uncertain about the relative merits of the treatments, we still have plenty of situations where different groups of researchers genuinely prefer different experimental treatments. In such a case personal equipoise is not in conflict with the exploration of different options. That said, personal equipoise does prevent situations where individual physicians or researchers knowingly enroll patients in treatments that they themselves deem unpromising. In cases where very few researchers prefer a therapy, it may not be ethically possible to promote the ideal exploration of different possibilities.

4.2 Improving Existing Communities

To this point we have considered a set of ideal solutions to the bandit problem–greedy strategies–and evaluated the practicalities of structuring our scientific communities to approximate the ideal solutions. Now, we will shift our perspectives a little bit to look at real epistemic communities that in many ways exhibit non-ideal characteristics. Drawing on the modeling literature in section 2, we

[10] Kitcher (1990) and Strevens (2003) present models suggesting that sometimes this worry about unfairness might not apply. In particular, they argue that credit incentives promote cognitive division of labor (i.e., incentivize researchers to pick less promising projects) because fewer individuals are working on these. Thus there is already a greater chance of being the one who will get credit for discovery in less popular areas of research. However, even if credit incentives naturally lead to a beneficial division of labor, this does not necessarily mean they lead to the best levels of division of labor.

instead ask: where are real scientific communities now, and how do we make them better? I.e., what are practical, implementable changes that can improve the status quo? In particular, we focus on realms where there are fairly clear failures impeding scientific progress.

There is a growing empirical literature showing that even though recruitment programs have brought diverse practitioners into research communities, their expertise, testimony, and epistemic output are not always properly recognized. For instance, Settles et al. (2019) conducted interviews at a predominantly white institution and found that faculty of color experience epistemic exclusion, characterized by a devaluation of their research topics, methodologies, etc. Moreover, Deo (2019, 47) presents a study finding that most women law professors experience "silencing, harassment, mansplaining, hepeating, and gender bias." Simply put, our real epistemic communities are not ones in which everyone's status as knowers is equally recognized and credited. Marginalized scholars experience epistemic silencing (Dotson, 2011).

As mentioned before, Wu (2021b) finds that this kind of epistemic silencing and exclusion can be detrimental to the community as a whole, but create epistemic advantages for the marginalized group. In epistemic communities with a history of epistemic silencing, then, the entire group might benefit from measures intended to strengthen the voices of marginalized members. This is one practical way to bring the benefits of transient diversity from one sub-group to the larger epistemic community. Notice that this mechanism for improvement is different from those proposed in previous work focusing on the ways that demographic diversity can lead to diversity of practice. Background experiences of the world resulting from personal identity can, indeed, lead to this sort of diversity, but here we are focused on how status as member of a marginalized group simpliciter can create epistemic advantages.

As noted, industry interests can influence science in a number of ways. For instance, industry scientists may choose to only publish and promote scientific research that adheres to certain goals (e.g. the Tobacco Strategy discussed in Oreskes and Conway (2011) and Weatherall et al. (2020)). Especially relevant here is the fact that industry scientists may not share evidence that they gather and discoveries they make, though they may still read and learn from research published by other scientists. Academic researchers adhere to the "communist norm" that all research should be shared (Merton, 1942, 1979; Strevens, 2017; Heesen, 2017). But industry research is often proprietary.

This practice of industry scientists creates conditions for transient diversity to arise. Wu (2021a) develops models similar to her ones exploring testimonial injustice, but where one group refuses to share, rather than one group refusing to listen. When a small group of scientists refuses to share their evidence, but the rest of the community follows the communist norm, the small group tends to develop true beliefs at a higher rate. This is because the rest of the community receives less evidence on average, and thus spends more time exploring undesirable options. The result is a transient diversity of practice that only benefits industry scientists.

In this scenario, one might think that the benefits of transient diversity that industry scientists gain are extremely unfair to the rest of the epistemic community, especially since the mechanism (not sharing) that gives rise to the transient diversity is in violation of a central norm that scientists are expected to adhere to.[11] It is hard to know just how this issue might be addressed. One possibility would be to legally obligate industry to share proprietary research. This, of course, might disincentivize industrial groups from ever performing said research. Research is costly, and industry is often willing to pay those costs in order to gain knowledge that others do not have. Solutions might require sharing of research after some set period of time, or else depend on patenting to protect industry in a way that incentivizes funding research without also keeping this research private.

4.3 Mechanisms for Transient Diversity

We now turn to the last main part of this section: assessing the various mechanisms that give rise to transient diversity in network models, as surveyed in Section 2. These models fall into two broad categories. The first offer proposals for possible interventions that we can (or even, should) implement in epistemic communities, in order to gain transient diversity that communities would not otherwise have. In short, they offer policy recommendations to achieve transient diversity. The second set of models identify possible mechanisms which might already promote transient diversity in real communities, but do not necessarily offer policy proposals.

This distinction is important because it will determine how we think about these different causes for transient diversity. For models that offer policy recommendations, we might want to then consider whether the proposed interventions are actionable, practical, ethical, and whether the benefits that we would potentially gain from the interventions outweigh the harms. On the other hand, for models that address mechanisms for existing transient diversity, we might want to instead focus on whether these mechanisms are present in target epistemic communities, whether other factors may have interfered with the mechanism, and whether the benefits of transient diversity are worth maintaining given potential harms. Of course, some of the considerations remain the same regardless. For instance, in both cases, we assess whether the benefits of transient diversity are worthwhile given epistemic and ethical harms associated with it.

With this in mind, let us further discuss the models from Section 2. Both Zollman (2007, 2010) and Kummerfeld and Zollman (2015) intend their models to either offer solutions to the problem of (lack of) diversity or offer policy recommendations through institutional involvement.

Kummerfeld and Zollman (2015)'s proposal of incentivizing individual scientists to continue testing suboptimal theories at a small rate comes the closest to approximating the ideal solutions in decision theory. As we have discussed earlier, for many research communities it will not be practical for individual

[11] See Bright and Heesen (2021) for an argument that to be scientific is to adhere to the communist norm.

scientists to switch methodologies regularly. So a Kummerfeld and Zollman (2015)-like proposal would be more implementable in some epistemic communities if we instead incentivize a small subgroup of scientists to explore suboptimal theories, with the hope that these incentives are also strong enough to curb the free rider problem. Barring practical constraints discussed earlier, Kummerfeld and Zollman (2015)'s proposal may be effective at generating the benefits of transient diversity without too much epistemic and ethical cost. In general, their results and others suggest that there may be real benefits to grant giving agencies that come from the diversity of practice when they focus on exploratory research and unpopular theories.

What about Zollman (2010)'s proposal of limiting communication among scientists? First there are some practical constraints to implementing such a proposal. Scientists tend to want to communicate their research, and have many venues for doing so. It is unclear how a community would go about slowing this communication. Perhaps professional agencies could host fewer conferences, journals could publish more slowly, grant giving agencies could cut funding for meetings, travel and talks etc.

It is unclear whether the benefits of transient diversity under this proposal outweigh the potential harms, but we suspect they often will not. As mentioned in Section 3, Rosenstock et al. (2017) shows that when the learning problem is fairly easy, the benefits of transient diversity are small. This means that unless one is very clear about the sort of learning situation scientists are in, the proposal in question runs the risk of slowing down learning with little benefit. Furthermore, if communications become too limited, communities run the risk of polarization (a la Wu (2021b)). Moreover, limiting communications among scientists goes against the communist norm which, as noted, is a central norm of scientific research. Given these potential harms, it seems that there are better ways to ensure transient diversity of practice than artificially limiting communication between scientists.

This said, there are cases where it may be worthwhile to temporarily limit communication in order to improve discovery. To offer an example of this (though bandit models do not perfectly apply) we think that the limited communications among the four imaging teams in the Event Horizon Telescope project may be epistemically beneficial. In this case, the four teams work in isolation. They each use a different method to develop imaging algorithms, train their algorithms against test data sets, and finally, produce their own images of the black hole from real data collected from multiple observatories around the world, before convening to compare images (Galison and Newman, 2021). The four teams end up producing highly similar images while working in isolation. Recently, sets of labs have been teaming up for "many labs" papers, which run independent tests of the same hypothesis before combining results at the end of the project (Ebersole et al., 2016; Klein et al., 2018). In both cases, independent tests can help determine whether results are replicable before they are published and begin influencing the community. Note, though, that in both of these cases communication is limited between groups who are interested in the same theories, but who want to preserve diversity in their tests. The benefits of limited communication here are related to, but somewhat different from, those identified using bandit models.

What about the other mechanisms that lead to transient diveristy, such as testimonial injustice (Wu, 2021b) and confirmation bias (Gabriel and O'Connor, 2021)? In both cases we have pretty good reasons to believe that these factors are widely present in many epistemic communities. Also, in both cases these mechanisms seem to be poor candidates for policy recommendations, since few individual scientists would want to knowingly commit injustices or engage in reasoning biases.[12] Our question now is: are the benefits of transient diversity worthwhile to keep in these cases? Of course, there may be no choice, since these factors depend on deep facts about human psychology, but we take it to be worthwhile, nonetheless, to discuss whether or not attempts to eliminate (or preserve) them are right headed.

The answer seem clear in the case of testimonial injustice. The benefits of transient diversity come with a number of epistemic and ethical harms. Epistemic injustice directly harms individuals as knowers and community members. Relatedly, for marginalized groups, although they may garner epistemic advantages in the sense of learning true beliefs more often and faster, they may not receive credit proportional to their epistemic achievements (Rubin, 2021). This is because the epistemic achievements possessed by the marginalized group may not be acknowledged and recognized by the wider epistemic community. Lastly, this mechanism may lead to polarization in the epistemic community, where the marginalized and dominant groups stably converge to different theories. Polarization often is associated with community dysfunction. Epistemic injustice, polarization, and credit deficit are all significant harms that outweigh any benefits of transient diversity in this case.

That said, as with any kind of systemic oppression, we should not expect testimonial injustice to go away easily in a short period of time, even with active efforts to decrease it. As a result, in many epistemic communities, we may continue to see subgroups of marginalized individuals with certain epistemic advantages. In cases like these, the rest of the epistemic community may have opportunities to gain epistemic insights from these groups. The suggestion here is to implement interventions designed at reducing testimonial injustice to improve the flow of ideas that have already benefited from transient diversity of practice.

The case with confirmation bias is considerably different. To start, while the tendency to update on evidence that confirms one's preferred theory may violate some norms of good inquiry, it does not seem to commit glaring injustice to others. Moreover, as discussed, low levels of confirmation bias facilitate the discovery of true belief by slowing down the community learning process. Given this, we think that in some cases, confirmation bias may bring surprising epistemic benefits. On this picture, active attempts to increase informational literacy by decreasing confirmation bias may sometimes have negative effects.

[12] Not to mention that in the case of testimonial injustice, the agents committing testimonial injustice learn worse on average in terms of accuracy and speed (Wu, 2021b).

This said, it seems risky to actually attempt to promote confirmation bias (or other, similar forms of irrational stubbornness) in science. There are real historical cases where scientific communities have polarized over matters of fact, to the apparent detriment of inquiry. O'Connor and Weatherall (2018) give a case study regarding chronic Lyme disease where this sort of polarization has arguably harmed a wide swathe of patients. And in any case, it is difficult to intervene on deep seated psychological biases. For these reasons, this does not seem the best lever for promoting beneficial levels of transient diversity in science.

4.4 Complex Problems and Transient Diversity

We have now concluded the main discussion of the paper, but want to address a serious limitation before concluding. To this point, we have relied heavily on the multi-armed bandit model of scientific exploration, both in outlining the benefits and harms of transient diversity, and also in assessing various mechanisms for promoting transient diversity. But, as noted in previous sections, this is not the only model of scientific exploration. And when we consider other models, this shifts the analysis. Of particular interest here are the NK-landscape models briefly described in section 2. As noted, in this sort of complicated problem space optimal solutions may not be easily accessible from all starting places. And, in particular, there are often local optima such that individuals who reach them must then explore less successful options before they can discover global optima.

There are many areas of science with similar structures, i.e., where adopting the best theory/option at one point in inquiry closes off other pathways that might lead to better options later. In such cases a much more radical level of exploration and diversity is merited than would be appropriate for bandit-model-type scientific problems. The mandate is no longer to explore apparently suboptimal options because they themselves might, in fact, be better than previously thought. The mandate is to keep exploring these options, even when it is clear that they are, indeed, suboptimal, because it is possible that they will lead to other, important discoveries.

Alternatively, there are simpler problems that demand less transient diversity. Bandit models generally assume that arms are independent, and thus that learning about one does not yield information about the others. In science, though, theories are often interrelated so that tests of topic A also inform topic B. In an extreme case, we can imagine situations where there are tests which can decide between two competing theories. If so, or if scientists face a particularly easy bandit problem, transient diversity of practice might not be particularly important (Rosenstock et al., 2017).

We might ask: is there some sort of way to know what the problem space of a scientific discipline looks like? If so, then maybe the goal could be to promote a more modest level of diversity in those areas where bandit-problems and similar models apply, and more radical diversity of practice in those areas where NK-landscape models apply. This does not strike us as a promising plan, though. In many cases, if some topic is on the cutting edge of scientific research, its structure, as a problem, is not well understood. Even problems that seem almost tailored to bandit models may have more complex structures. Take an example used by Zollman (2010). Suppose there is a well-understood drug A, and a new, experimental drug B. The goal is to figure out which drug has more efficacy for some condition. Tests, here, involve prescribing the two drugs and seeing how patients react. This is well-modeled by a bandit problem. Suppose that B turns out to be the worse option, and physicians begin to stop prescribing and studying it. But also suppose that B, when combined with some different therapy C, is tremendously beneficial. A problem that looked like a bandit problem thus turns out to be more complex.

This complicates recommendations for scientific communities tremendously. It may mean that the level and persistence of transient diversity necessary to ensure good epistemic progress is more dramatic than previously imagined. It also might mean that the explore/exploit trade-off inherent to transient diversity of practice is more serious. This raises questions such as: is it acceptable to keep using clearly unpromising medical therapies on the chance that they will generate future discoveries? Equipoise mandates reduced diversity of practice, despite its epistemic benefits, because what matters to human wellness is success right now.

This observation about NK-landscapes relates to another complaint one might make about using bandit models in this investigation. Typical models of transient diversity that employ bandits assume a well-defined, well-understood set of options for explorations. I.e., scientists can try A or B (or C or D). But the model does not include the generation of new theories. The landscape model does a better job of representing this process, as it often involves actors searching a space to look for new theoretical possibilities. Some of these investigations have suggested that cognitive diversity is important in that it might prompt individuals to test different theories, and explore different areas of theory-space (Thoma, 2015; Pöyhönen, 2017).[13] This claim complements arguments mentioned earlier from feminist philosophy of science, standpoint epistemology, and science studies that cognitive diversity is important in shaping choices of research questions and hypotheses generated (Okruhlik, 1994; Haraway, 2013).

This is all to say that there is a different and important sort of diversity of thought/practice from what we have been focusing on. Bandit models help us ask: what is the optimal distribution of investigation over a set of possibilities? But we also want to ask: how do we encourage diversity of thought and diversity of practice that leads to the exploration of new and unexpected possibilities? In this latter vein, proposals about diversifying the demographics of scientific communities seem promising. The idea is that demographically diverse community members may entertain a wide set of hypotheses, have different background assumptions, and use different inferential standards and practices. For instance, in the avian fire-spreading example we started with, local Indige-

[13] See also Weisberg and Muldoon (2009), though their original modeling work has been convincingly criticized by a number of authors (Thoma, 2015; Alexander et al., 2015; Pöyhönen, 2017; Pinto and Pinto, 2018).

nous residents are able to uncover a novel variable in fire management, due in large part to their intimate relationships with local environment and their unique generational knowledge keeping practices. Furthermore, the observations in this section lend further support to the proposal that funding sources, and publication venues, should actively try to promote at least some level of unusual, exploratory, and unpopular topics of investigation.

5 Conclusion

Several promising pathways have emerged from the discussion here. First, centralized bodies are important to coordinate research across a community and support exploratory research in particular. Second, promoting diverse voices is a relatively low-cost, low-risk way to increase exploration across topics in science and to improve benefits from transient diversity of practice. Third, network models may provide guides as to where to locate benefits of transient diversity in existing epistemic communities, e.g. marginalized subgroups and proprietary industry scientists.

With regards to the first proposal there has been increasing support of modified lottery funding in science. Reviewers tend to be drawn to proposals that are highly promising, safe, and familiar. This means that risky, exploratory, unusual, and unpopular topics tend to be rejected. A lottery ensures that at least some topics which seem less promising in the short term still receive funding. Typical proposals along these lines first reject grant applications that are clearly below the bar, maybe accept the most exceptional proposals, and then use a lottery to determine further funding. There is a further benefit to using lotteries, which arises in cases where there is genuine confusion and disagreement about how promising some theories are compared to others. As noted, this kind of disagreement is common in epistemic communities. It also complicates attempts by funding bodies to ensure ideal levels of exploratory research. Although lotteries cannot ensure ideal levels either, they do provide a way to increase diversity of practice even when different stakeholders in a community disagree about what this sort of diversity should look like.

In this paper, we have discussed a number of proposals and possibilities for ensuring a good level of transient diversity of practice in science, drawing on recent modeling work. Scientific communities, and scientific problems, are complex. This means that our discussion is necessarily tentative. Though the models we discussed may not apply to all scientific communities, our hope is to draw readers to this small body of literature that may inform and complement the larger discussion on social and cognitive diversity in science.

Acknowledgements

Many thanks to the members of our NSF group meeting for discussion and feedback on early drafts of this proposal—Clara Bradley, Matthew Coates, David

Freeborn, Yu Fu, Nathan Gabriel, Daniel Herrman, Aydin Mohseni, Ainsley Pullen, and Jim Weatherall.

References

Alexander, J. M., Himmelreich, J., and Thompson, C. (2015). Epistemic landscapes, optimal search, and the division of cognitive labor. *Philosophy of Science*, 82(3):424–453.

Berry, D. A. and Fristedt, B. (1985). Bandit problems: sequential allocation of experiments (monographs on statistics and applied probability). *London: Chapman and Hall*, 5(71-87):7–7.

Bonta, M., Gosford, R., Eussen, D., Ferguson, N., Loveless, E., and Witwer, M. (2017). Intentional fire-spreading by "firehawk" raptors in northern australia. *Journal of Ethnobiology*, 37(4):700–718.

Bright, L. K. and Heesen, R. (2021). To be scientific is to be communist.

Cor, K. and Sood, G. (2018). Propagation of error: Approving citations to problematic research.

Currie, A. (2019). Existential risk, creativity & well-adapted science. *Studies in History and Philosophy of Science Part A*, 76:39–48.

Deo, M. E. (2019). *Unequal profession: race and gender in legal academia*. Stanford University Press.

Derex, M. and Boyd, R. (2016). Partial connectivity increases cultural accumulation within groups. *Proceedings of the National Academy of Sciences*, 113(11):2982–2987.

Dotson, K. (2011). Tracking epistemic violence, tracking practices of silencing. *Hypatia*, 26(2):236–257.

Ebersole, C. R., Atherton, O. E., Belanger, A. L., Skulborstad, H. M., Allen, J. M., Banks, J. B., Baranski, E., Bernstein, M. J., Bonfiglio, D. B., Boucher, L., et al. (2016). Many labs 3: Evaluating participant pool quality across the academic semester via replication. *Journal of Experimental Social Psychology*, 67:68–82.

Fang, C., Lee, J., and Schilling, M. A. (2010). Balancing exploration and exploitation through structural design: The isolation of subgroups and organizational learning. *Organization Science*, 21(3):625–642.

Fehr, C. (2011). What is in it for me? the benefits of diversity in scientific communities. In *Feminist epistemology and philosophy of science*, pages 133–155. Springer.

Freedman, B. (2017). *Equipoise and the ethics of clinical research*. Routledge.

Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.

Fried, C. (1974). *Medical Experimentation Personal Integrity and Social Policy.*

Gabriel, N. and O'Connor, C. (2021). Can confirmation bias improve group learning?

Galison, P. and Newman, W. E. (2021). Interview with peter galison: On method. *Technology— Architecture+ Design*, 5(1):5–9.

Gittins, J., Glazebrook, K., and Weber, R. (1989). *Multi-armed bandit allocation indices*. John Wiley & Sons.

Gittins, J. C. (1979). Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(2):148–164.

Haraway, D. (2013). *Simians, cyborgs, and women: The reinvention of nature*. Routledge.

Haraway, D. J. (1989). *Primate visions: Gender, race, and nature in the world of modern science*. Psychology Press.

Heesen, R. (2017). Communism and the incentive to share in science. *Philosophy of Science*, 84(4):698–716.

Holman, B. and Bruner, J. (2017). Experimentation by industrial selection. *Philosophy of Science*, 84(5):1008–1019.

Jones, R. (2012). Fire-stick farming. *Fire Ecology*, 8(3):3–8.

Jönsson, M. L., Hahn, U., and Olsson, E. J. (2015). The kind of group you want to belong to: Effects of group structure on group accuracy. *Cognition*, 142:191–204.

Kitcher, P. (1990). The division of cognitive labor. *The journal of philosophy*, 87(1):5–22.

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams Jr, R. B., Alper, S., Aveyard, M., Axt, J. R., Babalola, M. T., Bahník, Š., et al. (2018). Many labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science*, 1(4):443–490.

Kuhn, T. S. (1977). Collective belief and scientific change. In *The essential tension*, pages 320–339. University of Chicago Press.

Kummerfeld, E. and Zollman, K. J. (2015). Conservatism and the scientific state of nature. *The British Journal for the Philosophy of Science*, 67(4):1057–1076.

Lai, T. L. and Robbins, H. (1985). Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics*, 6(1):4–22.

Lazer, D. and Friedman, A. (2007). The network structure of exploration and exploitation. *Administrative science quarterly*, 52(4):667–694.

London, A. J. (2009). Clinical equipoise: Foundational requirement or fundamental error. In Steinbock, B., editor, *The Oxford Handbook of Bioethics*. Oxford University Press.

Longino, H. E. (1990). *Science as social knowledge: Values and objectivity in scientific inquiry*. Princeton University Press Princeton.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization science*, 2(1):71–87.

Mason, W. and Watts, D. J. (2012). Collaborative learning in networks. *Proceedings of the National Academy of Sciences*, 109(3):764–769.

Mason, W. A., Jones, A., and Goldstone, R. L. (2008). Propagation of innovations in networked groups. *Journal of Experimental Psychology: General*, 137(3):422.

Mercier, H. and Sperber, D. (2017). *The enigma of reason*. Harvard University Press.

Merton, R. (1942). The normative structure of science. *Journal of Legal and Political Sociology*, (1):115–126. Originally titled "Science and Technology in a Democratic Order".

Merton, R. K. (1979). The normative structure of science.

Neale, A. V., Dailey, R. K., and Abrams, J. (2010). Analysis of citations to biomedical articles affected by scientific misconduct. *Science and engineering ethics*, 16(2):251–261.

Nersessian, N. J. (2019). Creating cognitive-cultural scaffolding in interdisciplinary research laboratories. *Beyond the Meme: Development and Structure in Cultural Evolution*, pages 64–94.

O'Connor, C. (2019). The natural selection of conservative science. *Studies in History and Philosophy of Science Part A*, 76:24–29.

O'Connor, C. and Weatherall, J. O. (2019). *The misinformation age*. Yale University Press.

Okruhlik, K. (1994). Gender and the biological sciences. *Canadian Journal of Philosophy*, 24(sup1):21–42.

Oreskes, N. and Conway, E. M. (2011). *Merchants of doubt: How a handful of scientists obscured the truth on issues from tobacco smoke to global warming*. Bloomsbury Publishing USA.

O'Connor, C. and Weatherall, J. O. (2018). Scientific polarization. *European Journal for Philosophy of Science*, 8(3):855–875.

Palmer, E. D. (1954). Investigation of the gastric mucosa spirochetes of the human. *Gastroenterology*, 27(2):218–220.

Pinto, M. F. and Pinto, D. F. (2018). Epistemic landscapes reloaded: An examination of agent-based models in social epistemology. *Historical Social Research/Historische Sozialforschung*, 43(1 (163):48–71.

Pöyhönen, S. (2017). Value of cognitive diversity in science. *Synthese*, 194(11):4519–4540.

Rosenstock, S., Bruner, J., and O'Connor, C. (2017). In epistemic networks, is less really more? *Philosophy of Science*, 84(2):234–252.

Rubin, H. (2021). Structural causes in citation gaps.

Russell-Smith, J., Cook, G. D., Cooke, P. M., Edwards, A. C., Lendrum, M., Meyer, C., and Whitehead, P. J. (2013). Managing fire regimes in north australian savannas: applying aboriginal approaches to contemporary global problems. *Frontiers in Ecology and the Environment*, 11(s1):e55–e63.

Settles, I. H., Buchanan, N. T., and Dotson, K. (2019). Scrutinized but not recognized:(in) visibility and hypervisibility experiences of faculty of color. *Journal of Vocational Behavior*, 113:62–74.

Smaldino, P. and O'Connor, C. (2021). Interdisiplinarity can aid the spread of better methods.

Stanford, P. K. (2019). Unconceived alternatives and conservatism in science: the impact of professionalization, peer-review, and big science. *Synthese*, 196(10):3915–3932.

Strevens, M. (2003). The role of the priority rule in science. *The Journal of philosophy*, 100(2):55–79.

Strevens, M. (2017). Scientific sharing: Communism and the social contract. *Scientific collaboration and collective knowledge*, pages 1–50.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Thoma, J. (2015). The epistemic division of labor revisited. *Philosophy of Science*, 82(3):454–472.

Venkatesh, B. and Goyal, S. (1998). Learning from neighbors. *Review of Economic Studies*, 65:595–621.

Vosoughi, S., Roy, D., and Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380):1146–1151.

Warren, J. R. and Marshall, B. (1983). Unidentified curved bacilli on gastric epithelium in active chronic gastritis. *The lancet*, 321(8336):1273–1275.

Weatherall, J. O., O'Connor, C., and Bruner, J. P. (2020). How to beat science and influence people: policymakers and propaganda in epistemic networks. *The British Journal for the Philosophy of Science*, 71(4):1157–1186.

Weisberg, M. and Muldoon, R. (2009). Epistemic landscapes and the division of cognitive labor. *Philosophy of science*, 76(2):225–252.

Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution.

Wu, J. (2021a). Communism and industrial science.

Wu, J. (2021b). Epistemic advantage on the margin: a network standpoint epistemology.

Zollman, K. J. (2007). The communication structure of epistemic communities. *Philosophy of science*, 74(5):574–587.

Zollman, K. J. (2010). The epistemic benefit of transient diversity. *Erkenntnis*, 72(1):17.