ORIGINAL RESEARCH

How should we promote transient diversity in science?

Jingyi Wu[1] · **Cailin O'Connor**[1]

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 an appropriate amount of diversity of practice in scientific communities?

Keywords Transient diversity · Epistemic diversity · Cognitive diversity · Social epistemology · Network modeling

1 Introduction

It has been widely argued that diversity of practice 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. What is more, instances of progress in science—from the Copernican revolution (Kuhn, 1977) to the move to include females in our theories of primate behavior (Haraway, 1989)—have involved periods during which scientists disagreed about which theories and approaches are best. For these

✉ Jingyi Wu
   Jingyi.Wu@uci.edu

✉ Cailin O'Connor
   cailino@uci.edu

1 Department of Logic and Philosophy of Science, University of California, Irvine, Irvine, CA, USA

reasons, many have argued that it is worth promoting and preserving a diversity of beliefs and practices within scientific communities.

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 also a wide ranging qualitative literature on this topic. 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 *practice* 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 (Fehr, 2011; Haraway, 1989; Longino, 1990). 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 Sect. 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 showing why transient diversity is important, and how to generate it. Section 3 considers a less popular topic—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. We draw on decision theory to clarify this point. In Sect. 4 we present the main contribution of the paper: an in-depth discussion of concrete proposals for maintaining appropriate levels of diversity in science, keeping in mind practical and ethical constraints. This section has three main parts. (1) We consider how scientific communities might mimic decision theoretic norms for exploration. That is, after identifying optimal levels of diversity and exploration in science we ask: is it possible to mimic these optimal levels in real communities? If so, how might we best approximate them? (2) We assess the usefulness of concrete mechanisms from the modeling literature for promoting transient diversity. And (3) we consider how the limitations of the modeling approach we use impacts our

---

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

discussion. Ultimately we argue that a number of mechanisms for transient diversity are not particularly promising avenues for future interventions because they are either impractical or unethical to implement. Three promising avenues involve (1) the use of funding bodies to coordinate research topics, (2) requirements for sharing industrial research, and (3) the promotion of work by epistemically marginalized scholars.[2] 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) argue for the importance of division of labor—where different community members tackle different problems—in science. As they point out, if all scientists work on the same problems important insights might be missed.[3] Smaldino and O'Connor (2023) 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 in science, 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 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 Bala and Goyal (1998) in economics. They 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 what appears to be the most successful action and reaping

---

[2] We argue for this point, though, for different reasons than those proposed in previous literature.

[3] 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.

its benefits. The arms here can represent practices in science that yield epistemic successes at different rates. For instance, the practices might involve treating ulcers with either antacids or antibiotics (Zollman, 2010), or either treating chronic Lyme patients with long term antibiotics or not (O'Connor & 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, actors 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 actors 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 his 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 practice in science. A community in these sorts of models can fail epistemically 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 with each other, are more likely to end up at the correct consensus.[4] In these less connected groups it is more likely that pockets of diverse beliefs are preserved, ensuring that each action is tested long enough to discover its

---

[4] The robustness of Zollman's results is challenged by Rosenstock et al. (2017) (see Sect. 3).

true properties. 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 less, 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.[5] 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 approaches to a problem, and preemptively settles on one. Lazer and Friedman (2007) and Fang et al. (2010) find that actors in less connected networks tend to find better solutions to these problems (though it takes them longer).[6]

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 preserve such diversity.[7] 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 Carnap (1952) and Kuhn (1977) that different inductive standards are acceptable for science, and work by Kitcher (1990) and Solomon (1992) and Solomon (2001) on the epistemic benefits of having stubborn individuals in a group. If some scientists are "irrationally" stubborn, this might (sur-

---

[5] Here N is a parameter for the dimension of the solution space, and K is a parameter for the complexity of the space. 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.

[6] 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, and combine their discoveries. See also Mason et al. (2008) and Jönsson et al. (2015) who yield similar findings. Derex et al. (2018) look at yet another sort of agent based model of cultural evolution and find that innovation is maximized with intermediate levels of connectivity. Mason and Watts (2012), on the other hand, find a general disadvantage to low connectivity in problem solving groups.

[7] However, perpetually intransigently biased actors may prevent the group from ever agreeing. See Holman and Bruner (2015).

prisingly) benefit the community since they insist on testing a diversity of possibilities rather than following only the most promising ones.[8]

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. In bandit models, groups with (mild) confirmation bias outperform groups without. 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.[9]

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 engage in this sort of exploration at a socially optimal level, yielding a free rider problem. They take this as a reason that institutions such as funding bodies and award-granting agencies might play a beneficial role in promoting exploratory science.

Wu (2022a) 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, but not vice versa due to power differentials. This might represent a case where members of one dominant racial or ethnic group ignore those in another, or members of one "mainstream" 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 in 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 out-group cannot.[10]

In related work, Wu (2022b) develops models similar to her ones exploring testimonial injustice, but where one group refuses to share, rather than one group refusing to listen. Industry scientists typically do not share evidence they gather or discoveries they make, though they still consume research published by other scientists. Academic researchers on the other hand adhere to the "communist norm" that all research should be shared (Heesen, 2017; Merton, 1942, 1979; Strevens, 2017). In Wu's models, when a small group of scientists refuses to share their evidence, that group tends to develop true beliefs at a higher rate. This is because the rest of the community receives less

---

[8] Related to this are arguments from Dang (2019) that the best collaborative groups hold different beliefs, because they avoid herding and spend more time justifying their beliefs.

[9] This finding fits with informal claims from Mercier and Sperber (2017) that confirmation bias might play a functional role in groups. See also Dorst (2021)'s argument that confirmation bias is a rational response in situations of ambiguous evidence.

[10] Wu (2022a) points out that her findings support a main thesis of standpoint epistemology, where a disadvantaged social status can lead to epistemic advantages.

evidence on average, and thus spends more time exploring undesirable options. The result is a transient diversity of practice that only benefits industry scientists.

Relatedly, Fazelpour and Steel (2022) build a multi-armed bandit model with two mutually distrusting subgroups. They find that a moderate level of distrust can actually improve group learning. Distrust creates a longer period of diversity of practice in the entire epistemic community.[11] They also consider a model with two subgroups where actors have the tendency to conform to in-group members, but the presence of out-group members dilutes this conformity. They find that heterogeneous groups learn better than homogeneous groups in the presence of mild conformity.[12] They argue from their models that demographic diversity might improve group performance, since inter-group distrust and intra-group conformity occur across many social identity groups [see e.g. Levine et al. (2014) and Phillips and Apfelbaum (2012)].[13]

It is important to note that across all these models, generally, factors which improve consensus on better theories also slow down the emergence of community consensus. This should make intuitive sense. If scientists explore more theories long enough to really get a sense of their merits, they will not swiftly settle on one theory. Zollman (2007, 2010) describes this as a trade-off between speed and precision in epistemic communities—by taking the time to explore, groups delay consensus but increase chances of getting it right. We return to related ideas in the next section.

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. These include (1) reduced network connectivity, (2) individual stubbornness (or different standards for induction), (3) confirmation bias, (4) the active promotion of exploratory or risky science by funding bodies, (5) inter-group distrust, (6) proprietary industry research, and even (7) testimonial injustice.

In the next section, we will turn to 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 can help 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

---

[11] A very high level of inter-group distrust, however, can harm group learning (Fazelpour & Steel, 2022). This is because such high level of inter-group distrust would generate very little inter-group communication, making it less likely for the entire community to converge during the period of rounds considered in the paper.

[12] Though see also Weatherall and O'Connor (2021) who find that conformity in homogeneous communities in bandit models is generally detrimental to group learning.

[13] It may be instructive to contrast their group-based mutual distrust with the belief-based mutual distrust employed in O'Connor and Weatherall (2018), where actors trust those with similar beliefs more. O'Connor and Weatherall (2018)'s simulations show that belief-based mutual distrust is harmful to group learning, and may even lead to stable polarization of beliefs. The group-based mutual distrust considered in Fazelpour and Steel (2022) is also different from mechanisms in Wu (2022a, b), which are asymmetric between subgroups.

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 germane 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 in bandit problems mentioned before. In order to find out about a bandit arm, actors have to pull it, and they have to pull it enough that they get a decent sample of outcomes. In order to learn about new possibilities in science, they likewise have to test them.[14] This means that to promote good learning, there must be periods of inefficiency. This raises a question for scientific communities, though: how can a group maximize the benefits of transient diversity while minimizing 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 for individuals engaged in bandit problems (Berry & Fristedt, 1985; Gittins, 1979; Gittins et al., 1989; Lai & Robbins, 1985). These solutions identify the optimal amount of exploration to ensure that 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 *greedy* strategies, for instance, that typically select the best option based on past observation but with some small probability explore other options (Sutton & Barto, 2018). Some of these strategies decrease the level of exploration over time. 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 actions.[15]

The last section highlighted a number of mechanisms for promoting transient diversity in science, but while all of these mechanisms promote exploration, none of them exactly tracks optimal exploratory strategies.[16] This means that while they each improve group performance on average, they do not necessarily minimize the harms of transient diversity.

Along these lines Rosenstock et al. (2017) argue that policy suggestions from Zollman (2007, 2010) to decrease communication would create massive inefficiencies. They point out that significant transient diversity is most necessary for difficult prob-

---

[14] Sometimes old tests are able to illuminate new theories as well as old ones. In Sect. 4.3 we will say a bit more about the sorts of problems in science that are not well modeled by bandit problems for this reason.

[15] Details will determine just which sorts of greedy strategies work best—how much time do researchers have? How much data can they gather? What are the costs to getting things wrong? We do not go into detail since, as will become clear in the next section, it is not possible to replicate normative solutions to these problems in real communities. Instead, the goal will be to approximate something close.

[16] Of course, these optimal strategies are designed for an individual to implement, not a community, but we could still imagine a community coordinating in such a way as to perform an equivalent strategy. More in the next section.

lems in science. (When the problem is easy, on the other hand, almost every community successfully solves it.) Then Zollman's proposal involves decreasing the flow of information 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 such inefficiencies as much as possible, while still exploring enough to yield good outcomes. In other words, we should try to figure out which mechanisms can promote just enough transient diversity, while avoiding its harms.

Section 4 focuses on how to do this given the lessons we have seen from models, and given features of real scientific communities. Before moving on to that discussion, though, we would like to highlight two further harms from transient diversity.

First, across different models the mechanisms that promote transient diversity also tend to lead to polarization. When these mechanisms slow learning too much they create situations where disagreement in the community becomes stable, rather than transient. This 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 only listen to the sets of evidence that fit with their current beliefs.[17] In the models from Wu (2022a) 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 Wu (2022b) too, the community may become polarized, with the scientists who share preferring the worse theory. In all these cases, there are harms that can arise from one group failing to ever adopt more successful practices. These models all track cases where too much diversity of practice, lasting for too long, is a bad thing.

Second, 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 the use of fossil fuels (O'Connor & Weatherall, 2019; Oreskes & Conway, 2011; 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 (Oreskes & Conway, 2011, pp. 17–22). Relatedly 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 hinder epistemic progress. In all these cases, diversity of beliefs and practice in science are weaponized to extend the period during which less-successful practices continue to be used beyond what

---

[17] 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.

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, we can ask: what are the best ways to promote transient diversity in science, given our understandings of the harms, benefits, and mechanisms for doing so? The goal of the rest of this paper will be to shed some light on this question. In moving forward, we should keep in mind a few considerations. (1) As described, we want to promote enough diversity to ensure good outcomes, while minimizing harms from testing suboptimal theories. (2) 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 that constrain how we might ensure transient diversity of beliefs.

While there are many ways to structure the discussion that follows, we will divide it into three main parts. We start with optimal solutions identified in the decision theory literature and ask in Sect. 4.1: in what ways can real scientific communities mimic these solutions? (Or not?) As we argue, there are a number of reasons an exact approximation to these solutions is difficult. But there are ways in which communities might move in the direction of such solutions, especially via the use of funding bodies to coordinate research strategies. In Sect. 4.2 we then move on to assess the usefulness of the various concrete mechanisms for promoting transient diversity introduced in Sect. 2. Here we argue for the use of funding bodies to coordinate research (again), the promotion of work by previously marginalized researchers, increasing demographic diversity in epistemic communities, and also requiring industry to share proprietary results. We wrap up the section in 4.3 by discussing limitations of the models used here, and what these limitations mean for policy proposals.

4.1 Towards 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 Sect. 3. 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 so that 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 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 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.

There are barriers to doing something like this in a real epistemic community. Most pressing is the fact that individual scientists make their own decisions about what topics to investigate (Strevens, 2013). 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, curiosity, etc. In most academic communities, central coordinating bodies cannot simply hand topics out to scientists and demand that they investigate them.[18] 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, can shape the overall exploratory tendencies of a community (Goldman, 1999; Kummerfeld & Zollman, 2015; Viola, 2015). In such a regime, most money could be devoted to the most promising theories, but smaller pots systematically devoted to less promising options.[19] The level of exploration promoted by such funding bodies might be sensitive to the sorts of problems and topics under exploration. When accuracy is very important, they could promote more exploration, as compared to cases where swift convergence on some decent theory is best. We think this proposal is a promising one. There are some further roadblocks, however, that need consideration.

Some disciplines are heavily funded by one, or just a few, centralized funding sources. In these disciplines, centralized sources have some ability to coordinate research across the community. In other disciplines, funding comes from a diverse set of sources and this sort of coordination will be more difficult to achieve. One solution in such disciplines might involve more explicit coordination between funding bodies. If one entity devotes its funds entirely to some topic of research, others might try to improve transient diversity by branching out.

One more difficulty with coordinating funding is the fact that industry funds a considerable portion of scientific research. In the US, for example, businesses perform 75% of research and experimental development (R&D) and fund 72% of it (Boroush & Guci, 2022). Industries may not share the same incentive structures as public funding agencies. So our current discussions may only be limited to public funding for academic science, and may not apply to situations where industry funding is involved. Below, though, we will discuss policy proposals aimed at requiring industry sharing.

---

[18] Furthermore there are both ethical and pragmatic reasons to preserve this sort of academic freedom. Fleisher (2018) and Dang and Bright (2021) argue that perhaps scientists should not focus so much on projects they most prefer in order to promote diversity of practice. But this may have negative effects on productivity and focus in science.

[19] 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 sort of science that looks relatively unpromising at the moment, but may still yield important discoveries (Currie, 2019; Stanford, 2019). This discussion stems in part from 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 & Zollman, 2015; O'Connor, 2019; Stanford, 2019). Moreover, funding bodies, like journal editors, may already be predisposed to select less epistemically diverse projects [see Heesen and Romeijn (2019)]. If so, then perhaps funding bodies might need to work extra hard to promote this sort of research.

If these are implemented, coordination between publicly-funded and industry-funded research may be more feasible.[20]

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 some 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 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 may be difficult to achieve when labs cannot just change gears to research entirely new topics, even if the funding is available to do so.

Some areas of research (such as computational social modeling) have relatively small start-up costs for switching research focus, and in those areas it should be easier to coordinate transient diversity via funding. Other areas of research (such as FMRI studies in cognitive science) may require expensive, specialized equipment, or lengthy retraining to switch projects (Goldman, 1999; Viola, 2015). Mitigating this issue is the fact that in a scientific community exploratory strategies can be distributed across a group. That is, 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.

There may be an ethical cost to this sort of solution. If small groups of researchers devote their time to less promising research, they take on extra risks. These researchers may be less likely to receive credit for discoveries (since discoveries are less likely in general). To promote the benefits of transient diversity fairly, then, it might be warranted to introduce external mechanisms to compensate exploratory scientists. In addition, such compensations should help promote exploration. This relates to proposals from Stanford (2019) and Currie (2019) about using funding and credit incentives to promote this sort of research.[21] (Of course, again, this sort of proposal is most warranted in cases where there are reasons to prefer accurate consensus over speed.)

Another difficulty arises vis a vis communication. In the decision theoretic heuristic, there is no need to communicate results, because one individual sees them all

---

[20] Thanks to an anonymous reviewer for raising this point. For more on industrial science and varieties of epistemic diversity, see Holman and Bruner (2017) and Wu (2022b).

[21] 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.

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.[22] 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 here, 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 do we efficiently divide the exploration of the underlying space? 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 in cases where decision makers can allocate labor across a community, it is sometimes hard to know how a decision maker ought to allocate labor given the complexities of real scientific evidence. If so, one cannot approximate optimal exploratory strategies.

One way to proceed might involve using lotteries to help with funding decisions. Lotteries can ensure that a diversity of projects are funded, without requiring central coordinating groups to facilitate exploration. Lotteries also help address the issue that reviewers tend to be drawn to proposals that are highly promising, safe, and familiar. This often means that risky, exploratory, unusual, and unpopular topics tend to be rejected, decreasing diversity of practice. Typical proposals for lottery funding in science 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.[23] In particular, weighted lotteries that tend to fund preferred approaches with higher probabilities, but also fund other approaches with lower probabilities, might work

---

[22] 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 (Cor & Sood, 2018; Neale et al., 2010).

[23] Note that Lee et al. (2020) argue that this lottery funding scheme with an initial threshold may perpetuate black-white funding gaps, because there exists black-white disparity in this preliminary evaluation stage.

well.[24] Again note that the use of lotteries will be more appropriate in some cases on our analysis. They will be most useful when (1) it is worthwhile for the community to take time exploring many options in order to yield accurate consensus and (2) there is legitimate disagreement about the promise of various options.

We will discuss one more ethical constraint related to the use of coordinating bodies to promote transient diversity. There are cases where tests themselves cause significant harms. Testing nuclear bombs, for instance, requires contaminating some area with radioactive material. Because of such harms diversity of practice is not always worth promoting even when it might lead to important epistemic progress.

There are less clear cases related to the discussion in Sect. 3. Testing unpromising therapies on humans for example 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. In cases where there is genuine lack of consensus across the scientific community, though, centralized funding bodies can help by assuring that different researchers test the different therapies they prefer. In cases where very few researchers prefer a therapy, of course, it may not be ethically possible (or desirable) to promote the epistemically ideal exploration of different options.

To summarize: exactly approximating decision theoretic solutions in the messy reality of a scientific community will not always be easy, or even possible. But we can improve transient diversity by using funding bodies as a way to coordinate research. Depending on the details of the scientific community at hand, it may be warranted for these funding bodies to coordinate with each other, to protect labs engaged with less-popular research topics, to specially compensate scientists who do risky research, to use lotteries in cases where there is no agreement about which theories are most promising, and to protect equipoise by distributing funds to researchers with diverse beliefs.

4.2 Mechanisms for transient diversity

We now turn to assessing the various mechanisms that give rise to transient diversity in network models, as surveyed in Sect. 2. These models fall into two broad categories. The first offers proposals for possible interventions for achieving transient diversity. The second identifies mechanisms for transient diversity that may already be present in real communities, but do not necessarily offer policy proposals.

This distinction is important because it affects how we think about these different mechanism 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

---

[24] There is a literature on whether/how to employ lotteries in funding decisions. We do not survey it here, but see for example Fang and Casadevall (2016), Gross and Bergstrom (2019), Smaldino et al. (2019), Adam (2019), and Avin (2019a, b).

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, and whether the benefits of transient diversity are worth maintaining given potential harms.

With this in mind, let us further discuss the models from Sect. 2. We will start with Zollman (2010), Kummerfeld and Zollman (2015), and Fazelpour and Steel (2022) who all offer policy recommendations of some sort.

Kummerfeld and Zollman (2015)'s proposal is to incentivize individual scientists to continue testing suboptimal theories at a small rate, perhaps using funding mechanisms. This comes the closest to approximating optimal solutions in decision theory. For the reasons discussed in Sect. 4.1, Kummerfeld and Zollman (2015)'s proposal may be effective at generating benefits of transient diversity without too much epistemic and ethical cost. In general, their results support the claim that in the right cases there may be real benefits when grant giving agencies promote exploratory research and unpopular theories. Of course, as discussed, due to diverse practical constraints, the strategies for doing so might look different in different disciplines.

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.

Generally, we think that the harms from such measures outweigh potential benefits. As mentioned, Rosenstock et al. (2017) show 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. In other words, this intervention risks an inefficient level of exploration as discussed in Sect. 3. Moreover, limiting communications among scientists goes against the communist norm which specifies that academic science should be shared as widely as possible. This norm plays an important role in the spread of new knowledge.[25] Given these potential harms, we think there are better ways to ensure transient diversity of practice than artificially limiting communication between scientists.

This said, there may be cases where it is worthwhile to temporarily limit communication in order to improve discovery by providing robustness checks and limiting the spread of error. For example, the four imaging teams in the Event Horizon Telescope project worked in isolation. They each used a different method to develop imaging algorithms, trained their algorithms against test data sets, and finally, produced their own images of the target black hole from real data before convening to compare images (Galison & Newman, 2021). By producing similar images in isolation, they increased their confidence in these images, in part because they limited the possibility

---

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

that social influence would spread mistakes between the groups.[26] Similarly, in "many labs" papers multiple groups run independent tests of the same hypothesis before comparing results at the end of the project (Ebersole et al., 2016; Klein et al., 2018). These independent tests can help determine whether results are replicable before they are published. Note 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 thus related to, but different from, those identified using bandit models, where the idea is to preserve diversity across the theories tested.

The last policy recommendation we consider is Fazelpour and Steel's proposal to increase demographic diversity. Their models show clear epistemic benefits of having heterogeneous subgroups in a community if there is already (a small amount of) distrust or conformity. Moreover, there does not seem to be much epistemic or ethical harm associated with increasing demographic diversity generally (and, as we will note later, there are many arguments for its benefits). However, we caution against (mis)interpreting Fazelpour and Steel's models as advocating for increasing, or preserving, inter-group distrust and conformity. There is a great deal of evidence suggesting too much conformity and too much distrust can lead to epistemic harms (Fazelpour & Steel, 2022; Flaxman et al., 2016; O'Connor & Weatherall, 2018; Pariser, 2011; Weatherall & O'Connor, 2021).

Before continuing, it is important to note that so far we have evaluated these policy recommendations from the perspective of promoting diversity of practice. However, one may think that in some epistemic communities, there is currently too much diversity, leading to polarization or inefficiency.[27] For instance, there is no mandate to continue promoting diversity of opinions related to whether there is an anthropogenic global warming.[28] In some of these cases, funding bodies might improve outcomes by decreasing funding for fringe research. In others, increasing communication between researchers through conferences or "white papers" may help promote consensus formation a la Zollman (2010).[29]

Now let us turn to mechanisms that lead to transient diversity, but that have not been associated with explicit policy proposals. These include testimonial injustice (Wu, 2022a), industrial proprietary science (Wu, 2022b), confirmation bias (Gabriel & O'Connor, 2021), and intransigence (Zollman, 2010). In these cases we have good reasons to believe that these factors are widely present in many epistemic communities. Also, in each case these mechanisms seem to be poor candidates for policy recommendations, since few individual scientists would want to knowingly commit injustices or engage in poor reasoning.[30] Our question now is: are these factors worth preserving for their benefits to transient diversity? Of course, there may be no choice, since these factors depend on deep facts about human psychology, but we take it to be

---

[26] See also Bright and Heesen (Forthcoming).

[27] Thanks to an anonymous reviewer for raising this point.

[28] Though there may still be disagreements about how to combat global warming.

[29] We do not consider attempts to decrease diversity of practice by decreasing demographic diversity for obvious ethical reasons.

[30] Not to mention that in the case of testimonial injustice, the actors committing testimonial injustice learn worse on average in terms of accuracy and speed (Wu, 2022a).

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 gained in this way 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, 2022). Lastly, this mechanism may promote too much transient diversity, leading to polarization, which is often associated with community dysfunction. Epistemic injustice, polarization, and credit deficit are all significant harms that outweigh benefits from transient diversity gained.

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. 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, p. 47) presents a study finding that most women law professors experience "silencing, harassment, mansplaining, hepeating, and gender bias."[31] Simply put, our real epistemic communities are not ones in which everyone's status as knowers is equally recognized and credited (Dotson, 2011). In such communities 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. Furthermore this is not a proposal that risks harm, but rather one that arguably promotes an ethical and epistemic good.[32]

Remember that Wu (2022b) outlines another interpretation for her models, where industry researchers benefit from transient diversity by refusing to share their own proprietary research. This creates a situation where private interests exploit public goods for their own benefit. One way to improve scientific progress, then, is to spread industry findings more widely. It is hard to know just how to go about this. One possibility is to legally obligate industry to share proprietary research. This, of course, might disincentivize industrial groups from 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. These solutions have the

---

[31] Furthermore, Dobbin and Kalev (2016) find that existing programs such as diversity trainings are ineffective at reducing bias. See Kinney and Bright (2021) for an argument based on risk-weighted expected utility on why diversity trainings are ineffective.

[32] 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 the status of belonging to an epistemically marginalized group simpliciter can create epistemic advantages.

advantage of financially and epistemically incentivizing industry to fund research, since the initial period of privatization means that they may develop better findings faster due to the transient diversity present. But the public and the rest of the scientific community would eventually benefit from industry's research epistemically as well.[33]

The case with confirmation bias and intransigence is considerably different. To start, while these tendencies may violate some norms of good inquiry, they do not seem to commit glaring injustice to others. Moreover, as discussed, low levels of confirmation bias/intransigence facilitate the discovery of true belief by slowing down the community learning process. On this picture, active attempts to increase informational literacy, by decreasing confirmation bias for example, may potentially have negative effects. This said, it seems risky to actually attempt to promote confirmation bias (or intransigence) in science. There are real historical cases where scientific communities have polarized over matters of fact, to the apparent detriment of inquiry.[34] Santana (2021) argues that despite potential benefits of intransigence, there are real costs as well. Stubborn scientists may harmfully impact public trust in science by countering consensus. And in any case, it is difficult to intervene on deep seated psychological traits. For these reasons, this does not seem the best lever for promoting beneficial levels of transient diversity in science.

To summarize: the models discussed in Sect. 2 yield several active policy proposals that we think promising. First, they again support the use of centralized bodies to promote exploratory or less promising research. Second, they suggest that we should promote the presence of demographically diverse researchers in science, and promote their work. Third, they support measures to require the sharing of industrial research.

4.3 Complex problems and transient diversity

We have now finished the main discussion of the paper, but want to address a 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 Sect. 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

---

[33] Note that here we focus on industry-funded in-house research because this kind of industry-funded research is more likely to be withheld from the rest of the scientific community. But presumably industry-funded academic research (or publicly-funded academic research) may be withheld too, for instance because researchers would self censor in order to get further funding. In these cases, having a policy to require sharing may be appropriate too.

[34] O'Connor and Weatherall (2018), for example, give a case study regarding chronic Lyme disease where this sort of polarization has arguably harmed a wide swathe of patients.

who reach them must then explore less successful options before they can discover global optima.[35]

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 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 have situations where there are tests which 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 it makes sense 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.

In many cases, though, 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 slowly 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 kind of case complicates recommendations for scientific communities. It may mean that the level and persistence of transient diversity necessary to ensure good epistemic progress is more dramatic. If so, the mechanisms we outline for promoting transient diversity might need to be amplified, or continued for a long time. Adding further complexity is the observation from Sect. 3 that industry interests can take advantage of transient diversity to confound public belief, in which case it might be better to promote swift consensus in science. Scientists and science policy makers will have to make their best judgments case by case in order to successfully promote transient diversity. Both empirical work and the models discussed in this paper can help them do so.

NK landscapes raise one last issue with bandit models in this investigation. Typical models of transient diversity that employ bandits assume a well-defined, well-

---

[35] See Wu (2022b) for why transient diversity generated via withholding evidence is less robust in NK landscape models than in bandit problems. See Wu (manuscript) for a distinctive mechanism of generating transient diversity in the NK landscape problem.

understood set of options for explorations. That is, scientists can try A or B (or C or D), and they know their options. But the model does not include the generation of new theories. Landscape models of scientific progress on the other hand involve actors searching a space to look for new theoretical possibilities. Many 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 (Pöyhönen, 2017; Thoma, 2015).[36] 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 (Haraway, 2013; Okruhlik, 1994).

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 wider set of hypotheses, have different background assumptions, and use different inferential standards and practices. Together with the various strategies to promote diversity of practice discussed above, the community may eventually benefit epistemically from cognitive diversity that stems from social diversity.

To summarize: It is not always easy to know the structure of a scientific problem space. Across different problem spaces, transient diversity is important, and the mechanisms we have identified can be used to promote it. This said, more or less transient diversity may be ideal. Furthermore, demographic diversity in a community may be beneficial in promoting the exploration of new or unusual hypotheses.

5 Conclusion

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.

Several promising pathways have emerged from the discussion here. First, centralized bodies are important to coordinate research across a community and support exploratory research where appropriate. Second, promoting the work of previously epistemically marginalized scholars and increasing demographic diversity is a relatively low-cost, low-risk way to improve benefits from transient diversity of practice

---

[36] See also Weisberg and Muldoon (2009), though their original modeling work has been criticized by a number of authors (Alexander et al., 2015; Fernández Pinto & Fernández Pinto, 2018; Pöyhönen, 2017; Thoma, 2015). All these researchers focus primarily on a simple, two-dimensional landscape model.

and to increase exploration across topics in science. Third, interventions to require sharing of industry science should benefit epistemic communities.

Declarations

Conflict of interest The authors declared that they have no conflict of interest.

References

Adam, D. (2019). Science funders gamble on grant lotteries. *Nature, 575*(7785), 574–575.
Alexander, J. M., Himmelreich, J., & Thompson, C. (2015). Epistemic landscapes, optimal search, and the division of cognitive labor. *Philosophy of Science, 82*(3), 424–453.
Avin, S. (2019a). Centralized funding and epistemic exploration. *The British Journal for the Philosophy of Science, 70*(3), 629–656.
Avin, S. (2019b). Mavericks and lotteries. *Studies in History and Philosophy of Science Part A, 76,* 13–23.
Bala, V., & Goyal, S. (1998). Learning from neighbors. *Review of Economic Studies, 65*, 595–621.
Berry, D. A., & Fristedt, B. (1985). *Bandit problems: Sequential allocation of experiments*. Monographs on statistics and applied probability (Vol. 5(71–87), p. 7). Chapman and Hall.
Boroush, M., & Guci, L. (2022). *Research and development: Us trends and international comparisons*. National Science Foundation.
Bright, L. K., & Heesen, R. (Forthcoming). To be scientific is to be communist. *Social Epistemology*.
Carnap, R. (1952). *The continuum of inductive methods* (Vol. 198). University of Chicago Press.
Cor, K., & Sood, G. (2018). Propagation of error: Approving citations to problematic research. http://kennethcor.com/wp-content/uploads/2018/08/error.pdf.
Currie, A. (2019). Existential risk, creativity & well-adapted science. *Studies in History and Philosophy of Science Part A, 76*, 39–48.
Dang, H. (2019). Do collaborators in science need to agree? *Philosophy of Science, 86*(5), 1029–1040.
Dang, H., & Bright, L. K. (2021). Scientific conclusions need not be accurate, justified, or believed by their authors. *Synthese, 199*(3), 8187–8203.
Deo, M. E. (2019). *Unequal profession: Race and gender in legal academia*. Stanford University Press.
Derex, M., & Boyd, R. (2016). Partial connectivity increases cultural accumulation within groups. *Proceedings of the National Academy of Sciences of the United States of America, 113*(11), 2982–2987.
Derex, M., Perreault, C., & Boyd, R. (2018). Divide and conquer: Intermediate levels of population fragmentation maximize cultural accumulation. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*(1743), 20170062.
Dobbin, F., & Kalev, A. (2016). Why diversity programs fail. *Harvard Business Review, 94*(7), 14.

Dorst, K. (2021). Rational polarization. SSRN 3918498.

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., Brown, E. R., Budiman, N. I., Cairoj, A. H., Capaldi, C. A., Chartier, C. R., Chung, J. M., Cicero, D. C., Coleman, J. A., Conway, J. G., ... Nosek, B. A. (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., & Schilling, M. A. (2010). Balancing exploration and exploitation through structural design: The isolation of subgroups and organizational learning. *Organization Science, 21*(3), 625–642.

Fang, F. C., & Casadevall, A. (2016). Research funding: The case for a modified lottery. *mBio.* https://doi.org/10.1128/mBio.00422-16

Fazelpour, S., & Steel, D. (2022). Diversity, trust, and conformity: A simulation study. *Philosophy of Science, 89*(2), 209–231.

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

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

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly, 80*(S1), 298–320.

Fleisher, W. (2018). Rational endorsement. *Philosophical Studies, 175*(10), 2649–2675.

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

Fried, C. (1974). *Medical experimentation personal integrity and social policy*. North Holland.

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

Galison, P., & Newman, W. E. (2021). Interview with Peter Galison: On method. *Technology| Architecture+ Design, 5*(1), 5–9.

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

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

Goldman, A. I. (1999). *Knowledge in a social world*. Oxford University Press.

Gross, K., & Bergstrom, C. T. (2019). Contest models highlight inherent inefficiencies of scientific funding competitions. *PLoS Biology, 17*(1), e3000065.

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.

Heesen, R., & Romeijn, J.-W. (2019). Epistemic diversity and editor decisions: A statistical Matthew effect. *Philosophers' Imprint*.

Holman, B., & Bruner, J. P. (2015). The problem of intransigently biased agents. *Philosophy of Science, 82*(5), 956–968.

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

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

Kinney, D., & Bright, L. K. (2021). Risk aversion and elite-group ignorance. *Philosophy and Phenomenological Research*. https://doi.org/10.1111/phpr.12837

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, R. B., Jr., 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* (pp. 320–339). University of Chicago Press.

Kummerfeld, E., & 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., & Robbins, H. (1985). Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics, 6*(1), 4–22.

Lazer, D., & Friedman, A. (2007). The network structure of exploration and exploitation. *Administrative Science Quarterly, 52*(4), 667–694.

Lee, C. J., Grant, S., & Erosheva, E. A. (2020). Alternative grant models might perpetuate black-white funding gaps. *The Lancet, 396*(10256), 955–956.

Levine, S. S., Apfelbaum, E. P., Bernard, M., Bartelt, V. L., Zajac, E. J., & Stark, D. (2014). Ethnic diversity deflates price bubbles. *Proceedings of the National Academy of Sciences of the United States of America, 111*(52), 18524–18529.

London, A. J. (2009). Clinical equipoise: Foundational requirement or fundamental error. In B. Steinbock (Ed.), *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.

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

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

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

Mercier, H., & 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* (pp. 267–278). The University of Chicago Press.

Neale, A. V., Dailey, R. K., & 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. In *Beyond the meme: Development and structure in cultural evolution* (pp. 64–94). University of Minnesota Press.

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., & Weatherall, J. O. (2018). Scientific polarization. *European Journal for Philosophy of Science, 8*(3), 855–875.

O'Connor, C., & 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., & 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.

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

Pariser, E. (2011). *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin.

Phillips, K. W., & Apfelbaum, E. P. (2012). Delusions of homogeneity? reinterpreting the effects of group diversity. In *Looking back, moving forward: A review of group and team-based research*. Emerald Group Publishing Limited.

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

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

Rubin, H. (2022). Structural causes of citation gaps. *Philosophical Studies, 179*, 1–23.

Santana, C. (2021). Let's not agree to disagree: The role of strategic disagreement in science. *Synthese, 198*(25), 6159–6177.

Settles, I. H., Buchanan, N. T., & 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., & O'Connor, C. (2023). Interdisiplinarity can aid the spread of better methods. *Collective Intelligence*. https://doi.org/10.1177/26339137221131816.

Smaldino, P. E., Turner, M. A., & Contreras Kallens, P. A. (2019). Open science and modified funding lotteries can impede the natural selection of bad science. *Royal Society Open Science, 6*(7), 190194.

Solomon, M. (1992). Scientific rationality and human reasoning. *Philosophy of Science, 59*(3), 439–455.

Solomon, M. (2001). *Social empiricism*. The MIT Press.

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. (2013). Herding and the quest for credit. *Journal of Economic Methodology, 20*(1), 19–34.

Strevens, M. (2017). Scientific sharing: Communism and the social contract. In T. Boyer-Kassem, C. Mayo-Wilson, & M. Weisberg (Eds.), *Scientific collaboration and collective knowledge* (pp. 1–50). Oxford University Press.

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

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

Viola, M. (2015). *Some remarks on the division of cognitive labor*. *RT. A Journal on Research Policy and Evaluation*. https://doi.org/10.13130/2282-5398/4768

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

Warren, J. R., & 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. (2021). Conformity in scientific networks. *Synthese, 198*(8), 7257–7278.

Weatherall, J. O., O'Connor, C., & 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., & 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. In D. F. Jones (Ed.), *Proceedings of the VI international congress of genetics*, Ithaca.

Wu, J. (2022a). Epistemic advantage on the margin: A network standpoint epistemology. *Philosophy and Phenomenological Research*. https://doi.org/10.1111/phpr.12895

Wu, J. (2022b). Withholding knowledge.

Wu, J. (manuscript). Better than best: Epistemic landscapes and diversity of practice in science.

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.