

DOI: 10.1017/psa.2023.129

This is a manuscript accepted for publication in *Philosophy of Science*.

This version may be subject to change during the production process.

Better than Best: Epistemic Landscapes and Diversity of Practice in Science

Jingyi Wu

Department of Philosophy, Logic and Scientific Method, London School of Economics.

*Corresponding author. Email: Jingyi.Wu@lse.ac.uk.

Abstract

When solving a complex problem in a group, should group members always choose the best available solution that they are aware of? In this paper, I build simulation models to show that, perhaps surprisingly, a group of agents who individually randomly follow a *better* available solution than their own can end up outperforming a group of agents who individually always follow the *best* available solution. This result has implications for the feminist philosophy of science and social epistemology.

1. Introduction

When solving a complex problem in a group, should group members always choose the best available solution that they are aware of? This question arises when there is a group of people coming together to figure something out. They may be solving a scientific problem, or generating social or cultural innovation. If they care about their epistemic success, should they always choose the epistemically most successful option at the moment?

In this paper, I build simulation models to show that, perhaps surprisingly, a group of agents who individually randomly follow a *better* available solution than their own can end up outperforming a group of agents who individually always follow the *best* available solution. The reason for this result relates to the concepts of transient diversity (Zollman 2010; Wu and O'Connor 2023; Smaldino et al. 2022) and cognitive division of labor (Kitcher 1990; Weisberg and Muldoon 2009; Thoma 2015) in epistemic communities. The “better” strategy preserves a diversity of practice in the community for some time, so the community can survey a range of solutions before settling down.¹ The “best” strategy, by contrast, may lock the group in a suboptimal position that prevents further exploration. In a slogan, “better” beats “best.”

My models are adapted from Lazer and Friedman (2007)’s model where a network of agents is tasked with solving a sophisticated epistemic landscape problem called the NK landscape problem. Here, agents search in a space with multiple “peaks.” They only know the solutions of their neighbors on the network, and their own results of (limited)

¹Though, the “better” strategy preserves a diversity of practice *only* when social learning is not too frequent (see §3).

local exploration, so they may fail to ever discover better solutions globally. Agents in the model face an exploration and exploitation trade-off: do they exploit the solution they currently have and the local peak nearby, or do they explore other regions of the landscape for possibly better solutions? In my models, the “better” strategy allows for a high degree of exploration within the community, even though at every single time step, every agent’s expected payoff is strictly no greater than what they would have gained with the “best” strategy.

This result is significant because first, it reveals a tension between individual and group decision-making. Here, groups learn better in the long run when their members do not always choose the best for themselves in the short run. This tension itself is not new in social epistemology. For instance, Mayo-Wilson et al. (2011) proposes the Independence Thesis, which states that individual and group rationality may come apart. My results demonstrate this thesis in another modeling paradigm (c.f. the bandit problem in Mayo-Wilson et al. (2011)).

Second, many feminist philosophers of science (Longino 1990; Fehr 2011) suggest that different social groups tend to adopt different approaches to problem-solving, which can be represented by different starting points on an epistemic landscape. The “better” strategy explored in this paper, then, would be a good way to preserve these diverse approaches. Some of the solutions brought by marginalized groups may not seem promising, perhaps due to a historical lack of resources, but they may nevertheless become epistemically significant after explorations.

The model described here also makes technical contributions to the modeling science literature. First, the modeling paradigm I use, the NK landscape model, is very under-explored in the philosophy of science.² But I think this model represents a different yet (I will argue) important type of scientific inquiry. Second, the “better” strategy I introduce here is a new mechanism for transient diversity (see, again, Wu and O’Connor (2023) and Smaldino et al. (2022)). Moreover, this strategy, unlike other mechanisms, never makes an individual worse off from one round to another, and yet the community still performs relatively well. This may be a more practical strategy for generating epistemically-beneficial diversity.

This paper is organized as follows. In §2, I first provide a general interpretation of epistemic landscape models in the context of scientific problem solving. I then introduce the details of my model, including the two behavioral rules: “better” and “best.” In §3, I present the main simulation results of this paper. In §4, I consider a variation of the model: a mixed community where some agents adopt the “better” strategy, while others adopt the “best.” In §5, I draw implications of the results in the social and cognitive diversity literature.

²With the exception of Wu (2022) and Alexander et al. (2015). Though, the NK landscape model, with or without a network structure, is more widely used in theoretical biology (Kauffman and Levin 1987; Kauffman and Weinberger 1989), cultural innovation (Lazer and Friedman 2007; Gomez and Lazer 2019; Barkoczi and Galesic 2016), and organizational design (Ethiraj and Levinthal 2004; Siggelkow and Levinthal 2003; Marengo et al. 2000).

2. The Model

2.1. Interpreting the Epistemic Landscape Model

A landscape model contains a large number of points with varying heights. In the context of scientific problem-solving, we can use a landscape model to represent a group of scientists coming together to solve a problem by trying different research approaches. Each approach is a point on the landscape and has a score associated with it, representing its “epistemic significance,” e.g. how truth-conducive or fruitful it is. I interpret research approaches on such an epistemic landscape broadly. They can have many components, including research questions, methods, skills, instruments, etc. (Thoma 2015).

Epistemic landscape models can represent important aspects of scientific problem solving—scientists constantly communicate with others in their community and decide whether they should stick with the research approach they currently have or try new ones, either by exploring on their own or adopting an approach from the community. Epistemic landscape models, especially lower-dimensional ones with one or two peaks, have been used in the philosophy of science literature to model scientific problem-solving (Hong and Page 2004; Weisberg and Muldoon 2009; Thoma 2015).

2.2. The NK Landscape Model

The NK landscape model is a sophisticated multi-dimensional landscape with multiple peaks. This model was originally developed in theoretical biology to study how different variants of a gene work together to produce fitness (Kauffman and Levin 1987; Kauffman and Weinberger 1989). The solution space is N -dimensional, with binary strings (consisting of 0s and 1s) of length N as its points. For instance, if $N = 3$, then 001 is a point on the landscape, so is 101, 010, etc. Then, an algorithm with parameter K ($0 < K < N - 1$) is used to assign a score between 0 and 1 to each point.³ In the context of scientific problem-solving, we can think of each of these dimensions as a component of a research approach (research questions, tools, skills, etc.).⁴ The score of a research approach then depends on how well different components work together in synergy.

Roughly speaking, the parameter K in the NK landscape model determines how rugged the landscape is and how correlated “nearby” scores are. As K increases, the landscape becomes increasingly difficult to search. Though it is impossible to sketch a higher dimensional space, Figure 1 provides a stylized representation of the solution space as we vary K .⁵ When $K = 0$, the landscape is smooth with one single peak, and when $K = N - 1$, the landscape is totally chaotic, with the value of every point totally uncorrelated with nearby points. The most interesting space is when $0 < K < N - 1$. The landscape becomes rugged with multiple peaks, with nearby points somewhat correlated with each other. We will focus on this parameter space for the rest of the paper, since we can use it to represent complex scientific problems for which similar research approaches are somewhat correlated in epistemic significance. In this regime, it is often

³An online-only appendix describing the algorithm is available at <https://osf.io/jk6ed>.

⁴That each dimension only has binary options is an idealization. Investigating whether the results hold when this idealization is relaxed is an interesting follow-up project.

⁵This figure represents a landscape where $N > 2$, though it looks like a one-dimensional landscape.

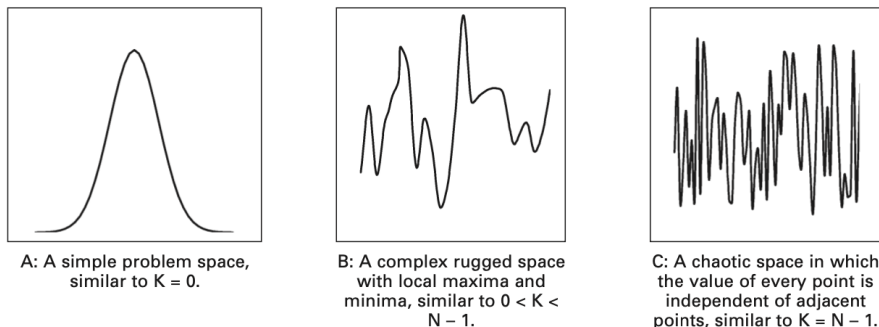


Figure 1. Stylized Representation of the Solution Space (Lazer and Friedman 2007)..

the case that a high-scoring solution is only accessible when exploring from a limited patch of the landscape.

Since the NK landscape model contains multiple peaks and a large number of solutions,⁶ it contains the following exploration and exploitation trade-off: do agents exploit the epistemic significance of their current solution (and the local optimum nearby), or do they keep exploring the landscape in hope of finding better solutions? If exploration is not enough, then agents may be stuck in local optima, without discovering more promising solutions. If exploration is too much, then agents waste time wandering around, trying out potentially low-scoring options.

2.3. Network Structures and Initialization

Having specified the solution space of the NK landscape problem, we now turn to the network structure of the model. At the start of each simulation, I have two separate communities of 100 agents, connected to each other via a directed Erdős-Rényi random network. This means that for every agent, the probability that it would form a link with another agent is a fixed number p in $[0, 1]$, and the links formed are not necessarily bidirectional. Every agent is assigned a starting solution, in the form of a binary string of length N . To facilitate direct comparisons, the two communities have identical initial conditions, including the solution space, starting solutions for individual agents, and network structure.

2.4. Behavioral Rules

Next, I introduce the two behavioral rules, “best” and “better.” All agents in one community follow the “best” behavioral rule, and all agents in the other community follow “better.” The only difference between the two behavioral rules is that, during social learning, agents with the “best” rule choose the best solution they are aware of, and agents with the “better” rule randomly choose a better solution than their own.

According to the “best” behavioral rule, in every V round ($1 \leq V \leq 5$), every agent chooses the best-performing solution among all their neighbors’ solutions to copy (if there are multiple with the same highest score, they randomly select one of the highest

⁶In the $N = 20$ case considered here, there are over a million solutions.

to follow).⁷ If they themselves have the best score among their neighbors, then they conduct a *local search* to try to improve their score. This means that they randomly choose a bit in their solution to change (1 to 0, and 0 to 1), and if the change brings a higher score, they switch to that solution.⁸ Otherwise, they maintain their current solution. In other rounds (rounds indivisible by V), they do a local search to try to improve their score. We call V the frequency of social learning. This behavioral rule is the same as the one explored in Lazer and Friedman (2007).

According to the “better” behavioral rule, every V round, every agent randomly chooses a better-performing neighbor and copy their solution.⁹ If they themselves have the best score among their neighbors, then they do a local search to try to improve their score. In other rounds, they do a local search.

In the context of scientific problem-solving, we can think of social learning as when scientists improve their own approaches by learning from their friends and collaborators. The “best” behavioral rule requires scientists to always adopt the best solution during social learning, and the “better” behavioral rule requires scientists to randomly choose a better one to mimic. Local search, on the other hand, is when scientists try to improve their own research approach by making small adjustments to it. Because it is typically the case that a high-scoring solution is only accessible from a limited patch of the landscape, agents can only discover their local peak when conducting local search, and social learning is the main mechanism for agents to move to other regions of the landscape.

3. Results

I run the model for long enough so that communities stabilize in their solutions. I set $N = 20$ and vary K , V , and p . I run 1,000 simulations for each parameter combination and present the average results.

The main result is that the community with the “better” behavioral rule ends up having a higher score than the community with the “best” behavioral rule (Figures 2 and 3). However, the “best” community performs better at the beginning of a simulation, and in general is faster at converging to a consensus solution.¹⁰ This is because a community with the “best” behavioral rule would quickly converge to the vicinity of the most promising solutions that they are aware of, while the “better” community would explore a variety of decent options and their close-by local peaks before building consensus, generating a diversity of practice. It takes the “better” community longer to survey the landscape, but its members are less likely to be stuck in low-scoring peaks.

This trade-off between speed and accuracy in social learning has previously been explored in models about how network connection impacts learning (Lazer and

⁷Each simulation starts at round 1, so agents do not go through social learning immediately unless $V = 1$.

⁸As a follow-up, one may be interested in models where agents alter more than one bit during local search. This may represent particular creative or resourceful individuals..

⁹Since an agent chooses a better performing *neighbor* to follow, if there are multiple neighbors with the same solution, the probability that this solution will be selected is proportionally increased. This is different from randomly choosing a better performing *solution* to follow.

¹⁰Here I only present agents’ eventual epistemic significance, their average performance per round, and their speed of convergence. For space reasons, I do not investigate other measures, e.g., an agent’s average epistemic significance across rounds (Pöyhönen 2017).

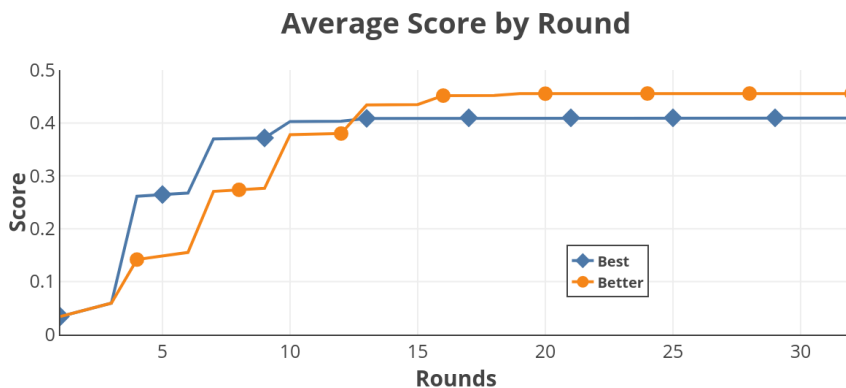


Figure 2. $N = 20$, $K = 10$, $V = 3$, $p = .5$. Both communities are stable after 30 rounds..

Friedman 2007; Zollman 2007; Zollman 2010). These models (from two different paradigms) show that a more sparsely connected community learns more accurately but more slowly, precisely because the community experiences a transient period during which a diversity of options is being tested. A more connected community is more likely to settle for an inferior option too early.

The present model introduces more texture to these previous results, since in a sense, the “better” behavioral rule seems to be an effective strategy to counter-balance the dangers of too much connection. As Figure 3 shows, the learning accuracy of the “best” community drops significantly as the network becomes more connected. But the “better” community maintains a high level of learning accuracy even as the community becomes more connected. Since limiting connectivity has been criticized as an impracticable way of improving epistemically-beneficial diversity (Rosenstock et al. 2017), encouraging individual community members to adopt something that mimics the “better” behavioral rule may be a plausible alternative.

While the result of “better” beating “best” is fairly robust overall, it is less robust when social learning happens every round.¹¹ To see why, let us consider a simplified case. Suppose that both the “better” and the “best” communities have A , B , and C as their starting solutions. Further, suppose that solution D is a local peak accessible from local search at B , and E is a local peak accessible from local search at C . Finally, suppose that $A < B < C < D < E$ in epistemic significance. In this scenario, the “better” community would be split between B and C first, and in cases where the agents with B solution discover D before agents with C solution discover E , the community may converge to D , without ever discovering E . In the “best” community, however, agents would all quickly converge to C first, and with sufficient local exploration, would discover E . When social learning slows down, this scenario happens less often, because agents in the “better” community would have enough “time” to explore the vicinity around both B and C sufficiently, so it is more likely that both D and E are discovered before the next social learning. This is a simplified case, but qualitatively similar situations happen with non-negligible probability in the full complex model. This suggests that in order for a diversity of practice to be beneficial to social learning, it has to be sustained in

¹¹The result holds more than 74% across all parameter combinations, and 100% when $V > 1$.

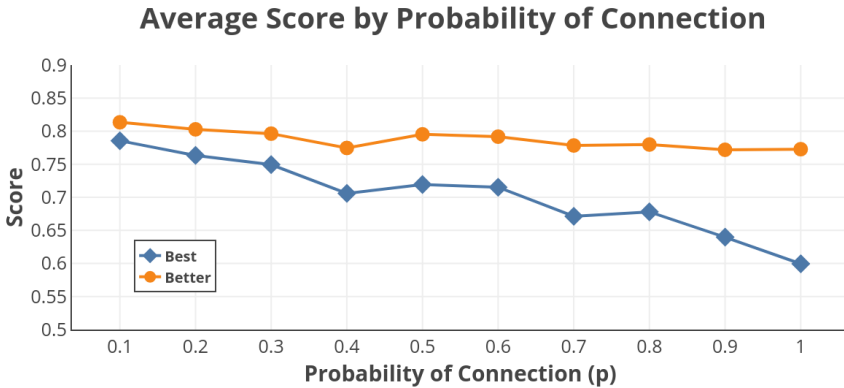


Figure 3. $N = 20$, $K = 5$, $V = 5$. Scores shown are average final scores of communities..

the community for some time to allow for sufficient local exploration; infrequent social learning makes this possible.

4. Variation: Mixed Community

I now introduce a variation of the model: a mixed community where some agents adopt the “better” strategy, while others adopt the “best.” In this community, even though all agents eventually converge to the same solution, agents who adopt the “best” solution reap more epistemic benefits at the beginning, as compared to agents who adopt “better” (Figure 4).

Comparing this community with the two communities studied in §3, we see that the mixed community ends up outperforming the “best” community, and scoring worse than the “better” community (Figure 4). The mixed community outperforms the “best” community precisely because of the agents who adopt the “better” strategy, and that creates a transient diversity of approaches in the community. The agents that adopt the “best” strategy here are essentially free-riding on the epistemic benefits that the “better” strategists provide. In so doing, they get the best of both worlds—they do (relatively) well eventually, while adopting high-scoring solutions in early game.

This creates a dilemma: the “better” strategists are useful to have in the community, but those agents may not be epistemically incentivized to keep their strategy. How should we encourage this epistemically exploratory behavior that benefits the community? I think promising solutions involve structuring the community in such a way that some agents find the “better” strategy attractive for other (intrinsic or extrinsic) reasons. For instance, Nguyen (2022) argues that *intellectual playfulness*—a disposition to try out new ideas for fun—functions as an intellectual “insurance policy” against what he calls *epistemic traps*. This, in a sense, is in line with my results—if some scientists are intrinsically motivated to try out exploratory solutions for the fun of it, then having them in an epistemic community and learning from them can be epistemically beneficial to the

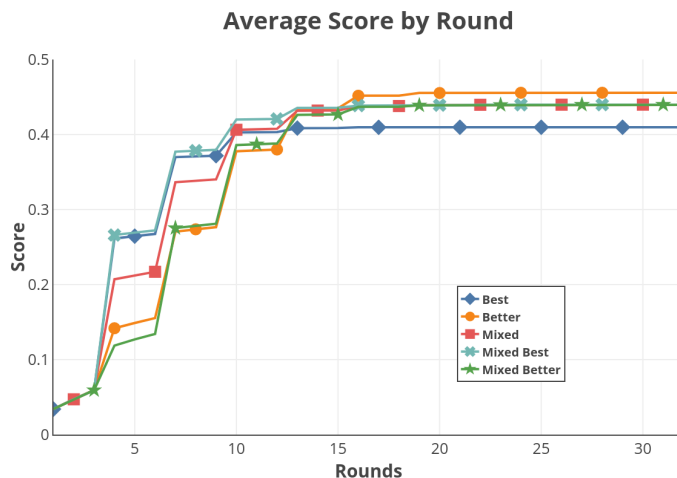


Figure 4. $N = 20$, $K = 10$, $V = 3$, proportion of the “better” group=.4. All communities are stable after 30 rounds. “Mixed” denotes the average score of the mixed community; “mixed best” denotes the average score of the “best” strategists in the mixed community; and “mixed better” denotes the average score of the “better” strategists in the mixed community..

community.¹² Another idea is to offer extrinsic incentives for exploratory ideas, such as coordinating funding agencies such that some amount of exploratory work is always funded and promoted. Moreover, various authors argue that even without dedicated funding, credit considerations alone may incentivize scientists to pursue exploratory research, since fewer individuals work on these topics (Kitcher 1990; Strevens 2003).

5. Coda: Social and Cognitive Diversity

In this paper, I present an epistemic landscape model in which a group of agents who randomly choose a better solution than their own can outperform a group of agents who always choose the best available solution. I argue that this result has a natural interpretation in the context of scientific problem-solving. A group of scientists who entertain a diverse range of reasonable research approaches for some time can outperform a group of scientists who always choose the best available research approaches. Before I close the paper, I will draw some implications of these results in the social and cognitive diversity literature.

First of all, note that the epistemically-beneficial diversity of practice in the main model is *not* the same as cognitive diversity. Cognitive diversity is usually understood as the presence of agents with different cognitive styles—different ways of gathering, processing, or acting on data (Hong and Page 2004; Pöyhönen 2017). The most epistemically successful community considered in this paper is a *homogeneous* group of agents

¹²While they share an exploratory attitude, an intellectually playful person completely disregards the truth value of the belief system they try out, but a “better” strategist still chooses a solution that is higher in epistemic significance. However, if some agents in the community choose to “randomly walk” in the landscape without any concern for epistemic significance, they will generate epistemically-beneficial diversity too.

who all follow the “better” rule.¹³ Moreover, even though a mixed community also performs well, it does well by virtue of having “better” strategists in the community, not by virtue of the *diversity* in cognitive styles.

Second, as briefly discussed in §3, in order for the diversity of practice to be epistemically beneficial in this model, the diverse range of solutions needs to be *sustained* in the community for some time. This allows for sufficient exploration of the local region around individual solutions, so local peaks are more likely to be discovered in between social learning. This means that if social learning is too frequent, it brings the same kind of epistemic harm as too much network connection (c.f. Lazer and Friedman (2007)). Infrequent social learning and sparse network structures can both ensure that a local region is sufficiently explored before an agent moves elsewhere.

Finally, feminist philosophers of science have written extensively about how diverse social groups tend to have diverse background beliefs and approaches to problem solving (Longino 1990; Fehr 2011), which may then be represented by diverse initial patches on an NK landscape model. If this is right, then the “better” behavioral rule is one way to help preserve initially plausible but not outstanding solutions from diverse (social) locations. Perhaps a solution from some marginalized social group does not stand out initially due to a historical lack of resources, but a stellar solution may be reachable if we explore in its vicinity. Moreover, I have also shown that having a diverse range of solutions simpliciter (more crudely—having diverse bodies in the room) is not enough, exploration in their vicinity needs to be supported in a sustaining fashion, so that agents do not prematurely switch to mainstream approaches without realizing the potential of their local perspectives.

Acknowledgments

Thanks to the NSF research group at UC Irvine’s Logic and Philosophy of Science Department, especially Yuhin Chung, Matthew Coates, David Freeborn, Nathan Gabriel, Ben Genta, Cailin O’Connor, and Jim Weatherall. Thanks to Hannah Rubin and Kevin Zollman for helpful comments. Thanks to Wybo Houkes and Hui Sun for suggesting literature on the NK landscape used in innovation science and organizational science. Thanks to audience members at the PSA 2022 in Pittsburgh, the workshop on agent-based modeling in epistemic communities in Bochum in 2023, and at the CLMPST 2023 in Buenos Aires. Most of the results presented here are replicated by Kevin Zollman in preparation for his forthcoming book. Thanks Kevin for this important work. This material is based upon work supported by the National Science Foundation under Grant No. 1922424. Simulation codes are available on <https://jingyiwu.org>.

¹³This is in line with Weisberg and Muldoon (2009)’s result that a homogeneous group of mavericks perform well.

References

- Alexander, J. M., J. Himmelreich, and C. Thompson (2015). "Epistemic landscapes, optimal search, and the division of cognitive labor". In: *Philosophy of Science* 82.3, pp. 424–453. DOI: <https://doi.org/10.1086/681766>.
- Barkoczi, D. and M. Galesic (2016). "Social learning strategies modify the effect of network structure on group performance". In: *Nature communications* 7.1, p. 13109. DOI: <https://doi.org/10.1038/ncomms13109>.
- Ethiraj, S. K. and D. Levinthal (2004). "Modularity and innovation in complex systems". In: *Management science* 50.2, pp. 159–173. DOI: <https://doi.org/10.1287/mnsc.1030.0145>.
- Fehr, C. (2011). "What is in it for me? The benefits of diversity in scientific communities". In: *Feminist epistemology and philosophy of science*. Springer, pp. 133–155. DOI: https://doi.org/10.1007/978-1-4020-6835-5_7.
- Gomez, C. and D. Lazer (2019). "Clustering knowledge and dispersing abilities enhances collective problem solving in a network". In: *Nature communications* 10.1, pp. 1–11. DOI: <https://doi.org/10.1038/s41467-019-12650-3>.
- Hong, L. and S. E. Page (2004). "Groups of diverse problem solvers can outperform groups of high-ability problem solvers". In: *Proceedings of the National Academy of Sciences* 101.46, pp. 16385–16389. DOI: <https://doi.org/10.1073/pnas.0403723101>.
- Kauffman, S. and S. Levin (1987). "Towards a general theory of adaptive walks on rugged landscapes". In: *Journal of theoretical Biology* 128.1, pp. 11–45. DOI: [https://doi.org/10.1016/S0022-5193\(87\)80029-2](https://doi.org/10.1016/S0022-5193(87)80029-2).
- Kauffman, S. and E. Weinberger (1989). "The NK model of rugged fitness landscapes and its application to maturation of the immune response". In: *Journal of theoretical biology* 141.2, pp. 211–245. DOI: [https://doi.org/10.1016/S0022-5193\(89\)80019-0](https://doi.org/10.1016/S0022-5193(89)80019-0).
- Kitcher, P. (1990). "The division of cognitive labor". In: *The journal of philosophy* 87.1, pp. 5–22. DOI: <https://doi.org/10.2307/2026796>.
- Lazer, D. and A. Friedman (2007). "The network structure of exploration and exploitation". In: *Administrative science quarterly* 52.4, pp. 667–694. DOI: <https://doi.org/10.2189/asqu.52.4.667>.
- Longino, H. E. (1990). *Science as social knowledge*. Princeton university press.
- Marengo, L. et al. (2000). "The structure of problem-solving knowledge and the structure of organizations". In: *Industrial and Corporate Change* 9.4, pp. 757–788. DOI: <https://doi.org/10.1093/icc/9.4.757>.
- Mayo-Wilson, C., K. J. Zollman, and D. Danks (2011). "The Independence Thesis: When Individual and Social Epistemology Diverge". In: *Philosophy of Science* 78.4, pp. 653–677. DOI: <https://doi.org/10.1086/661777>.
- Nguyen, C. T. (2022). "Playfulness Versus Epistemic Traps". In: *Social Virtue Epistemology*. Routledge, pp. 269–290.
- Pöyhönen, S. (2017). "Value of cognitive diversity in science". In: *Synthese* 194.11, pp. 4519–4540. DOI: <https://doi.org/10.1007/s11229-016-1147-4>.
- Rosenstock, S., J. Bruner, and C. O'Connor (2017). "In epistemic networks, is less really more?" In: *Philosophy of Science* 84.2, pp. 234–252. DOI: <https://doi.org/10.1086/690717>.
- Siggelkow, N. and D. Levinthal (2003). "Temporarily divide to conquer: Centralized, decentralized, and reintegrated organizational approaches to exploration and adaptation". In: *Organization Science* 14.6, pp. 650–669. DOI: <https://doi.org/10.1287/orsc.14.6.650.24840>.
- Smaldino, P. E. et al. (2022). "Maintaining transient diversity is a general principle for improving collective problem solving". In: DOI: <https://doi.org/10.1177/17456916231180100>.
- Strevens, M. (2003). "The role of the priority rule in science". In: *The Journal of philosophy* 100.2, pp. 55–79. DOI: <https://doi.org/10.5840/jphil2003100224>.
- Thoma, J. (2015). "The epistemic division of labor revisited". In: *Philosophy of Science* 82.3, pp. 454–472. DOI: <https://doi.org/10.1086/681768>.
- Weisberg, M. and R. Muldoon (2009). "Epistemic landscapes and the division of cognitive labor". In: *Philosophy of science* 76.2, pp. 225–252. DOI: <https://doi.org/10.1086/644786>.
- Wu, J. (2022). "Withholding Knowledge". In:

- Wu, J. and C. O'Connor (2023). "How should we promote transient diversity in science?" In: *Synthese* 201.2, p. 37. DOI: <https://doi.org/10.1007/s11229-023-04037-1>.
- Zollman, K. J. (2007). "The communication structure of epistemic communities". In: *Philosophy of science* 74.5, pp. 574–587. DOI: <https://doi.org/10.1086/525605>.
- (2010). "The epistemic benefit of transient diversity". In: *Erkenntnis* 72.1, p. 17. DOI: <https://doi.org/10.1007/s10670-009-9194-6>.