

Modeling epistemic communities

Draft, 5.10.2016

Samuli Pöyhönen & Jaakko Kuorikoski, forthcoming in M. Fricker, P.J. Graham, D. Henderson, N. Pedersen, J. Wyatt (eds.) *The Routledge Handbook of Social Epistemology*.

Word count: 5221

1. Introduction

Finding solutions to genuinely important epistemic challenges typically exceeds the capabilities of a single knower. Science, research and development laboratories, and the work of expert committees are all instances of knowledge production, which require coordinated effort from several agents. Furthermore, these situations essentially involve interaction and a division of cognitive labor among the members of the group or community: Difficult problems are attacked by dividing them into more tractable sub-problems, which are then allocated to subgroups and ultimately to individual group members. In this chapter, we use the notion of epistemic community to refer to such a group of agents faced with a shared epistemic task.¹ We regard division of labor between the members of the group as a necessary property of an epistemic community, so as to distinguish such groups from mere statistical or aggregative epistemic collectives, where there is no communication or coordination between group members (cf. Surowiecki 2005; Sunstein 2006). However, as we will see, the division of labor and the associated conception of cognitive diversity can be understood in several different ways.

¹ Following a usage common in the modeling literature, we understand an epistemic community as a social system consisting of producers of knowledge. In a broader sense of the notion (cf. Longino 1990), epistemic communities can be seen to encompass also the consumer side, i.e. those consulting the results.

In this chapter, we provide a review of the modeling work, which has aimed at understanding the functioning of epistemic communities.² Modeling done by social epistemologists resides at an interesting junction of various strands of inquiry. Many of the themes touched upon by social epistemologists (e.g., scientific reputation and authority, reward schemes in science) had already been studied by sociologists of science such as Robert Merton (1973) and Pierre Bourdieu (1975). Furthermore, the disciplinary origins of the various modeling methodologies used in social epistemology can typically be traced back to economics and decision theory, organization science, AI, and ecology. However, within social epistemology, formal modeling work has often formed its own niche and the integration of modeling work with the rest of research in social epistemology has often been less than satisfying. Hence, it is often not clear how the simple models of epistemic communities contribute to (a) the more general problems studied by social epistemology as a whole, and (b) how they should be connected to relevant findings from disciplines such as social psychology and organization studies.

After providing an overview of some of the most prominent modeling approaches in social epistemology (sections 2 to 6), in the final section of our review we try to go some way towards answering these questions. We sketch a general approach to interpreting models, thereby suggesting how they could be integrated with conceptual and empirical work (e.g., case studies) done in the rest of the field. Furthermore, by showing that the models in social epistemology are in many ways parallel to those developed in other fields that study collective knowing and problem solving, our contribution hopefully helps to position the philosophical work within this multi-disciplinary research area. We suggest that there is a clear epistemic benefit from seeing these parallels more clearly: It allows us to see where the strengths and blind spots of models presented in various fields are, and where promising future contributions might lie.³

² For alternative ways of organizing the material, see Weisberg 2010 and Muldoon 2013.

³ In this chapter, most of the attention is on epistemic communities in science. One might object to this by pointing out that the scope of social epistemology is broader than science. We focus on scientific problem solving for the following reasons: First, science has been the main target of much of the work that we review.

2. The invisible hand in science: From individual irrationality to collective rationality

In a now classic paper; Philip Kitcher (1990) set out to examine the division of cognitive labor in a community of scientists. By employing a combination of modeling tools from microeconomic theory (individual maximization of expected utility and equilibrium), Kitcher investigates how research resources should be allocated among alternative competing research programs, methods, or theories. In light of several examples drawn from the history of science, Kitcher argues that in many cases, a community of scientists should hedge its bets. All research effort should not be allocated only to the study of the currently most strongly confirmed or most promising alternative so as not to prematurely rule out potentially true but not yet well-confirmed hypotheses. Nevertheless, from the point of view of individual epistemic rationality, every scientist should pursue exactly the most promising avenue of research best supported by the currently available evidence.

There is, therefore, a discontinuity between the requirements of individual and collective rationality. Whereas an epistemic community might benefit from diversity provided by some stubbornness or biased appraisals of evidence, the rational behavior for each truth-motivated individual agent appears to be to join the best supported research program, method, or theory. This suggests that the rationality of individual agents (understood as each one optimizing their individual pursuit of the truth) is not sufficient for achieving good collective outcomes in research.

Kitcher's model aims to show that individual rationality is not necessary for collective epistemic efficiency either. Like Bourdieu (1975), Kitcher treats scientists as self-interested entrepreneurs in the search for personal prestige. The model suggests that under an appropriate reward allocation scheme, the individual incentives of a population of 'sullied' self-interested agents,

Secondly, scientific problem solving is a particularly challenging example: it is often not routine-like, nor is the social process of research hierarchically organized. Instead, the questions addressed by scientists are often open-ended, and the process is largely self-organized.

motivated not by truth as such but by individual glory garnered from finding the truth, drive the population towards the collectively optimal resource allocation.

Suppose that there are two alternative methods for finding out the structure of a Very Important Molecule. Truth, once discovered, is easy to recognize, and the probability of arriving at the truth with a given method is an increasing function of the number of people applying the method. Furthermore, suppose that when the truth is found by using a given method, each individual having used that method has an equal probability of being the one making the discovery and thus getting all the credit. Because of the way in which credit is allocated, choosing between the two alternative methods becomes a strategic decision involving not just the expected epistemic utility of the methods, but also the number of people already using them. Under these conditions, egoistic credit-seeking behavior may help to ensure that some resources are also allocated to the currently less well supported alternatives and thus to maintain crucial epistemic diversity in the community.

In the last chapter of his 1993 book *Advancement of Science*, Kitcher broadens his economics-inspired investigation of the epistemically pure and sullied agents by employing analytical machinery from population biology. This allows him to address questions about whether scientists should cooperate or go solo, how attribution of epistemic authority is done, and further questions related to trust, replication, and influence of scientific tradition on theory choice. We do not go into these arguments in detail here. Many of the topics introduced by Kitcher have been discussed in the models described in the sections below. Generally, the discrepancy between the micro and the macro has perhaps been the most lasting result of Kitcher's contribution to social epistemology, as it underscores the importance of studying the social processes of knowledge production (cf. Mayo-Wilson et al. 2011). Hence, epistemic communities can be studied as systems manifesting macro-level properties not reducible to the properties of their members, and the efficiency of such systems appears to be determined by at least the three following kinds of factors:

- The distribution of the cognitive properties of individual agents in the community (cognitive diversity)
- The organizational properties of the community (e.g., its communication structure, reward scheme)

- The nature and difficulty of the problem-solving task faced by the community

As we will see, the key concepts of cognitive diversity, division of cognitive labor, and resource allocation have subsequently been given several interpretations - often leading to confusion and difficulties in combining and comparing results from several models. We now turn to some of the extensions and examination of Kitcher's modeling approach.

3. Extending Kitcher's approach

Kitcher's argument rested crucially on the priority rule for apportioning credit. A further development of the resource allocation view is Michael Strevens' (2003) model of the role of reward schemes in resource allocation.

Already Merton (1973) had pointed out that priority disputes have always been a feature of modern science. There are no second prizes in research, only the first person to a discovery gets the recognition and prestige and virtually nothing is left for the runner-ups. But why is credit allocated due to this priority rule, and not based on a scientist's contribution to a research program (MARGE rule), or simply the success of her research program (GOAL rule)? And why is the priority rule applied in a peculiarly rigid way, even when the time difference between discoveries is a matter of days or hours?

Strevens argues that what is special about science as a collective endeavor is that the runner-up's contribution adds nothing to the collective good. Once a result has been discovered, no value to the collective is produced by discovering it again. Strevens uses basic economic principles ("what happens at the margin when the reward scheme changes") to compare the priority rule, MARGE, and GOAL, and shows that the priority rule allocates the comparatively largest proportion resources to the most potential research program: When reward is based neither on work or effort, or even on achievement as such, but on achieving-first, the resulting distribution of research effort over competing alternatives most closely resembles the collectively optimal allocation.

Note, however, that there is a problematic functionalist step in Strevens' explanation: without an additional mechanism, the mere fact of congruence between the individual and the collective good

fails to explain why the priority rule *in fact* prevails in actual scientific research. Furthermore, Piimies and Sappinen (mimeo) point out that the analytical approach taken by Kitcher (and Strevens) bears a striking resemblance to a branch of economics called neoclassical economics of innovation. A comparison to this literature reveals that most of Kitcher's model-based arguments are in fact highly sensitive to the selection of initial assumptions. Moreover, Kitcher exacerbates the problem by introducing seemingly ad-hoc modifications to his models "on the fly" - a move strictly forbidden by the modeling methodology in economics. For example, the central result about the connection between credit seeking and beneficial diversity rests crucially on the assumption that the agents are risk-neutral. Such fragility of results with respect to simplifying assumptions dramatically complicates the evaluation of the empirical implications of the models. This does not, by any means, invalidate Kitcher's (or his followers') achievements, but implies that such results should be considered more as theoretical suggestions, not as evidence or proof.

4. Theory choice in epistemic networks

Another key aspect of research modeled in Kitcher's investigations in the *Advancement* was a scientist's decision of whether or not to borrow results from others given their (observable) rate of hitting on the truth. Recently, the tension between individual theoretical exploration and the need to converge on a solution has been explored by Kevin Zollman (2010, 2013). Using ideas from statistics, economics, and network theory, Zollman models theory choice in terms of so-called bandit problems (see below).⁴ Zollman does not study epistemically sullied credit-seeking agents as Kitcher and Strevens do, but the same basic result holds: Individual rationality alone is not sufficient for reliable convergence to the truth on the social level.

Zollman portrays a scientist's choice between competing theories (or methods) as analogical to a learning problem faced by a gambler, who has to select between two slot machines with initially unknown winning rates. Each bandit is postulated to have some objective probability of success,

⁴ Zollman's model generalizes a framework proposed in economics by Bala and Goyal (1998).

and the payoff from winning from a machine is considered analogous to the payoff received from a successful application of a theory. Furthermore, pulling a machine amounts to getting more information about the truthfulness of a theory by running an experiment. However, because the agent's resources are limited, only a limited number of pulls can be conducted. Hence, she must devise a strategy for pulling the two levers so as to efficiently converge on the more lucrative machine.

The learning challenge presented by such bandit problems is non-trivial because of the trade-off between exploration and exploitation. In the short term, each agent gets the largest payoff by only running the best experiment so far discovered, whereas exploration of other alternatives is typically a precondition for long-term success. Zollman constrains his study to simple myopic agents who always sample only the best alternative and leave exploration to other members of their epistemic network. That is, in Zollman's simulations, in addition to individual experimentation, the agents also receive the experimental results obtained by their network neighbors. These sources of evidence are integrated with the agent's existing knowledge by Bayesian conditionalization.

The most well-known result from Zollman's models is that, perhaps counterintuitively, too much social information can be harmful. In densely connected epistemic networks, misleading evidence obtained in the early stages of the simulated research process can spread quickly through the network, suppress experimentation on alternative methods, and lead the whole community to converge on a suboptimal alternative. That is, it might be that correct hypotheses do not get a fair chance. For example, in the research on peptic ulcer disease, a single early study on the causes of the disease proved so influential in the field that the whole research community was wrongly convinced for a half a century that the illness could not be caused by bacteria (Zollman 2010).

However, densely connected networks perform comparatively better when agents' initial beliefs are strong. What these two findings together suggest is that what is common to epistemically successful communities is *transient diversity*, a proper balance between the diversity of beliefs and consensus. Such a balance allows, at first, sufficient exploration of alternatives, but eventually leads to convergence on truth.

Zollman and his coauthors have extended the bandit approach in subsequent papers (Mayo-Wilson et al. 2011; Zollman 2013). Many of Zollman's simplifying assumptions were also relaxed by Alexander (2013). He modifies Zollman's approach in three ways. First, instead of assuming Bayesian conditionalization in evidence uptake, he studies heuristic agents relying on reinforcement learning (cf. Russell and Norvig 2003, ch. 6). Secondly, the model does not assume a fixed structure of the epistemic network, but instead, network formation is assumed to be based on preferential attachment, where agents form connections to their peers based on the observed success of a peer's approach. Thirdly, Alexander further complicates the learning problem faced by the scientist agents by adding the possibility of theoretical innovation by attaching new arms to the bandit.

Alexander's model suggests that in this more difficult learning task, preferential attachment helps successful theories to spread in the population, and hence a population of connected agents can converge more quickly than individual learners. The spreading of a very good theory (of success probability close to 1) in the network is, however, conditional on the fact that either old theories get sometimes forgotten, or that agents time discount old information about theories. Both mechanisms prevent the network from getting locked in into suboptimal belief and communication patterns.

Compared to Kitcher's pioneering work, the simulation methods used by Zollman and Alexander make it possible to study several important properties of epistemic communities not within the reach of the earlier analytical models (such as the dynamics of belief change, heterogeneity, details of the learning heuristics used by agents, and the influence of the communication structure of the community on its epistemic performance). The big advantage of such agent based simulation models (ABMs) is that they allow viewing group problem solving not only as a formal decision problem, but rather as a form of coordinated social behavior arising from the cognitive mechanisms and interpersonal processes involved. Importantly, the models suggest that no one learning heuristic or communication structure of an epistemic community is optimal for all research tasks. This suggests that the computational study of epistemic communities should adopt the more fine-grained strategy of mapping the dependencies between the individual properties of agents, properties of the social system, and the task addressed by the community.

However, computational modeling also has its disadvantages. Compared to analytical models, simulations often come with more parameters, and it is often hard to see what really drives the obtained results. In addition to this opaqueness of the models, the results obtained often hold only in limited parts of the parameter space. For example, the generalizability of Zollman's results have been questioned both on theoretical and empirical grounds. Rosenstock et al. (forthcoming) demonstrate that the negative Zollman effect of connectivity is in fact highly sensitive to the selection of parameter values. Similarly, when Mason and Watts (2012) investigated experimentally the problem solving performance of epistemic networks, they found that highly connected networks of problem solvers outperformed sparser ones even in tasks thought to be conducive to phenomena similar to the Zollman effect.⁵

5. Opinion dynamics

Since the 1950s, social scientists have developed mathematical models addressing the question of why some social processes lead to polarization of opinions, whereas others lead to the formation of consensus. In the classical paradigm of opinion dynamics, agents have opinions represented by numerical quantities, and the model is used to investigate the way in which different initial profiles of opinions and different social processes of information exchange influence the convergence and clustering of opinions.

Hegselmann and Krause (2002, 2006, 2009) imported this modeling approach to social epistemology by adding to the model the effect of truth seeking: In addition to social information, some agents in the population also revise their opinions based on a signal coming from "the truth"

⁵ The problems with the Zollman model do not compromise the more general exploration-exploitation trade-off observed already by March 1991 and Lazer and Friedman 2007: A well-functioning epistemic community must strike a balance between exploration and exploitation. Whereas efficient dissemination and uptake of information between the nodes of a network improves the behavior of the epistemic community in the short term, allocating sufficient resources to exploratory activities is necessary for long term success.

(i.e., by running experiments). The two main analytical results from their model (HK) are (1) the funnel theorem and (2) the leading-the-pack theorem.

The first states that if all agents are truth seekers (even to a modest degree), the population will ultimately reach a consensus that lies arbitrarily close to the truth. The latter theorem concerns the more interesting situation in which only some of the agents aim at the truth while others rely solely on social information (Hegselmann and Krause's somewhat strained interpretation of the notion of division of cognitive labor). According to the theorem, for all initial profiles, even one agent with a truth signal suffices for the convergence of population on truth, on the condition that all non-truth seekers are connected⁶ to a truth seeker through a chain of other agents. The extensive set of simulations conducted by the authors examines the parameter spaces of variants of the model in order to assess the generality of the analytical results.

Hegselmann and Krause's results suggest that under appropriate forms of social exchange, even a small minority of truth-guided agents in an epistemic community can lead the whole community to the truth. Riegler and Douven (2009) have recently extended the HK model by allowing the transmitted information to be noisy, and by representing agents' belief states as sets of propositions in a finite language. Like the original HK model, Riegler and Douven's model cannot directly answer the question of how much weight we should assign to the opinions of our colleagues. The model seems, however, to refine Hegselmann and Krause's findings by suggesting conditions for the quality of evidence: only sufficiently good evidence transmitted between agents facilitates convergence to the truth, whereas poor quality evidence can lead to the fragmentation of the population into several disagreeing groups.

⁶ Connectedness here means that the opinions are not too far from each other and, consequently, communication between the two agents remains possible.

6. Diversity and collaboration

Although ostensibly about diversity – commonly associated with a division of cognitive labor and the accumulation of community results – most of the models above have surprisingly little to say about epistemic collaboration in research (D’Agostino 2009). Typically they depict a *competition* between individual agents which, under appropriate conditions, leads, as if by an invisible hand, to good collective outcomes. However, in science there is also more profound division of cognitive labor between agents, research groups, and even between scientific disciplines. The ability of a research community to solve problems that exceed the capacities of any individual actor or group is based on recursively dividing problems into smaller sub-problems solvable by finite cognitive actors, and subsequently piecing together the sub-solutions to answer the original complex question. Central aspects of cognitive diversity involved in such *distributed problem solving process* are not represented in any of the models above: there is no representation of (a) variation in research heuristics or (b) specialization on a topic or question, which would differentiate an agent from the others.

Both heterogeneous heuristics and specialization are aspects of what we call *cognitive diversity*: due to different backgrounds, training, and talent, researchers might have different reasoning styles and, hence, different agents could contribute differentially to various aspects of the research topic. Here we discuss two recent contributions to the study of cognitive diversity and its consequences to collective problem-solving.

Michael Weisberg and Ryan Muldoon (2009) put forward a model that depicts research in a scientific field as a population of agents foraging on an epistemic landscape. Different patches of the landscape correspond to different research approaches, and the height of a patch represents the epistemic significance of that approach. That is, within each scientific field, different combinations of questions, available background theories, and methods have different potential to uncover the significant truths in the domain of research.

What makes this a model of cognitive diversity is that, there are three different kinds of agents in the population: controls, followers and mavericks. Each kind of agent has a different way of doing research. Followers tend to adopt approaches already examined by others, mavericks avoid them, and controls are insensitive to social information and rely solely on their own information

about how to improve their epistemic lot. Weisberg and Muldoon's simulations suggest (once again) the intuitively plausible result that diversity is good for a scientific community. Their most surprising finding is, however, that the maverick social learning strategy which uses information about others to *avoid* doing the same as others is the most effective way to promote the progress of research on a scientific field.

By importing a fitness-landscape based modeling approach from ecology, Weisberg and Muldoon introduced a new way of implementing cognitive diversity in a simulation model. However, the reliability and generality of their results have been called into question. The problem is, yet again, in the non-robustness of the results. Critics have shown that replacing some of the assumptions made by Weisberg and Muldoon by other, just as plausible, ones, different and even contradictory results regarding the role of cognitive diversity can easily be derived (Alexander et al. 2015; Thoma 2015; Pöyhönen 2016).

An influential approach to modeling cognitive diversity, which has received attention in economics, organization science, and recently also in philosophy, was introduced by Lu Hong and Scott Page (2000, 2004). Building on the Newell and Simon's (1972) view of problem-solving as heuristic search, Hong and Page portray diversity as consisting of two components: each agent having her particular (a) way of representing the problem ('perspective') and a (b) set of rules for devising new problem solutions from earlier ones ('heuristic').

In Hong and Page (2000), the problem faced by a community of problem-solvers is represented as that of constructing a bitstring – a sequence of ones and zeros – associated with the highest possible payoff. The bitstring in effect represents a structured solution to a complex problem, each bit standing for a possible solution to a subproblem. Each agent is characterized by a set of heuristics for flipping the states of the bits in the string, and an agent's individual performance level is the expected value of the value function at the local optimum at which the individual's search is drawn to a halt. Collaborative group performance, in turn, is determined by the value of the configuration reached by having the agents sequentially apply their own heuristics to manipulate the same binary string. Once an agent cannot improve on a solution, the turn is given to the next agent employing a different bit-flipping rule.

In such a setting, the diverse perspectives and heuristics used by agents can lead to globally optimal collective solutions, although no individual agent can reach the optimum on her own. Hong and Page use the model to argue that in group problem solving, adding a new problem solver to the group often does not have a diminishing marginal utility for the collective: Depending on the order in which agents' heuristics are applied to the problem, it can even be so that the addition of the last agent has the largest marginal contribution to collective performance, if she happens to contribute the missing piece to the puzzle by applying a novel heuristic that helps the population to escape a local optimum.

Using a similar model, Hong and Page (2004) derived several additional results about the usefulness of cognitive diversity in problem solving. They show that under quite general conditions, a diverse group of randomly chosen agents can outperform a homogenous group consisting of individually best-performing problem solvers. They call this the diversity-trumps-ability theorem. The theorem suggests, roughly, that to maximize problem-solving efficiency, a company should rather hire 20 random problem solvers than 20 of the best performers, who, being the best and hence equally good, would have to be cognitively similar. The result also suggests a conjecture regarding science policy: Perhaps the current emphasis on individual "excellence" in science funding should be tempered with policies aimed at maintaining diversity?

7. Conclusion: What can philosophers learn from models?

So what, in the end, have models of collective knowing thus far achieved and what are the future prospects for the field? Whereas much, even most of, science has become fundamentally model-based, in philosophy modeling methodology is still relatively new. Consequently, work in the field is sometimes characterized by the lack of a clear understanding about the goals and the added value of modeling in general, and of common agreed-upon methodological constraints.

What makes a good model? First of all, worthwhile model results should not be immediately obvious logical consequences of the used assumptions. Models, both analytical and simulation, are externally supported arguments, and the model as such does not do any work, if seeing the assumptions already suffices to make the conclusions obvious (Kuorikoski and Pöyhönen, unpublished). With hindsight, some of the early analytical models of science suffer from such theoretical sterility (see, for example, Goldman and Shaked 1991). Second, a model should be

theoretically well motivated in that it should be built on clear theoretical ideas, from which further implications are drawn with the help of the model.⁷ In social epistemology, however, much of the modeling work appears to be driven by the modeling framework itself. Model templates are often borrowed from elsewhere and simply given a new (not necessarily very convincing) interpretation in terms of epistemic properties of social systems.

Thirdly, models should be reliable in the sense that the results depend only on the empirically interpretable substantial assumptions and do so in a way, which makes the empirical interpretation of the results as straightforward as possible. As we have already pointed out, many of the models in social epistemology are problematic in these respects, because the central results often only hold in a very restricted part of the parameter space or under very specific technical assumptions. If the key results are not robust with respect to essentially arbitrary modeling assumptions, which are made only for the sake of making the model “work,” and which lack a clear empirical interpretation, then inferring from the demonstrated results to real epistemic social systems amounts to little more than guesswork. Running models and simulations can be seen as a new exciting method for philosophers, but we have a lot to learn from scientists engaged in modeling.

Nevertheless, we firmly believe modeling to be a valuable methodology for thinking about epistemically well-designed social systems. Theorizing in natural language about “emergent” properties that arise from complex organization and interaction is extremely limiting and prone to error – especially since when faced with structurally and interactionally complex systems, our pre-theoretic intuitions often tend to become increasingly unreliable. Modeling is in many cases a prerequisite for having theoretical understanding about the dynamics of collective cognitive systems: convergence, equilibria, and the micro-mechanisms involved. Working with models is an invaluable part of theorizing and analysis, because it forces us to make ideas precise, and allows the transparent examination of their coherence. ABMs in particular make it possible to

⁷ An elegant example of such model building can be found in Hegselmann and Will (2013).

examine complex scenarios where our intuitions and reasoning falter. The combination of a large group of heterogeneous agents following simple rules often leads to results that are impossible to anticipate by using only analytical or “conceptual” methods.

One possible strategy for improving the abstract models in social epistemology would be to calibrate them with empirical data. However, relevant data about properties of research fields, problems, and research strategies employed by scientists is not readily available – and often it is not even clear how such evidence should be obtained. So while empirical calibration is a laudable aim, we are not there yet. However, we see the models reviewed here as serving a different purpose. Rather than being high-fidelity representations of particular target systems “out there,” they function as proof-of-concept exercises about the possible mechanisms underlying collective intelligence. They are theoretical arguments, not virtual experiments.

The most promising way forward in the modeling work in social epistemology is to aim at a better integration with the existing modeling traditions in neighboring fields. As things stand, the existing philosophical models are often isolated from empirical literature as well as from each other, and do not form a methodologically coherent progressive research program. However, the research literature on human problem solving, both in individual and social contexts, is extensive, and various research fields such as cognitive science, sociology of science, social psychology, and economics have contributed to explaining both the failures as well as the surprising strengths of problem-solving groups. Furthermore, there is a long-standing modeling tradition on collective heuristic problem solving in management studies and organization science. If we take Kitcher’s original definition of social epistemology seriously, there should not be any reason for social epistemologists to ignore such work, nor should there be any discontinuity between philosophical social epistemology and the empirical social science of group problem solving.

Acknowledgements

Thanks to David Henderson for his helpful comments and suggestions on the manuscript. The research for this chapter has been funded by The Academy of Finland and the University of Helsinki.

References

- Alexander, J. 2013. Preferential Attachment and the Search for Successful Theories. *Philosophy of Science*, 80:5, 769-782
- Alexander, J., Himmelreich J., Thompson C. 2015. Epistemic Landscapes, Optimal Search, and the Division of Cognitive Labor. *Philosophy of Science* 82:3,424–453
- Bala, V., and S. Goyal. 1998. Learning from neighbours. *Review of Economic Studies*, 65, 565–621.
- Bourdieu, P. 1975. The specificity of the scientific field and the social conditions of the progress of reason. *Social Science Information*, 14, 19-47,
- D’Agostino F (2009) From the organization to the division of cognitive labor. *Politics, Philosophy & Economics* 8:1,101–129
- Goldman, A. and M. Shaked 1991. An economic model of scientific activity and truth acquisition. *Philosophical Studies*, 63, 31–55
- Hegselmann, R. and U. Krause. 2002. Opinion dynamics and bounded confidence models, analysis and simulation. *Journal of Artificial Societies and Social Simulation*, 5:3.
- Hegselmann, R. and U. Krause. 2006. Truth and cognitive division of labour. First steps towards a computer aided social epistemology. *Journal of Artificial Societies and Social Simulation*, 9:3.
- Hegselmann, R. and U. Krause. 2009. Deliberative exchange, truth, and cognitive division of labour: A low-resolution modeling approach. *Episteme*, 6:2, 130–144.
- Hegselmann R. and O. Will. 2013. From Small Groups to Large Societies: How to Construct a Simulator? *Biological Theory*, 8:2, 185-194.
- Hong, L. and S. Page. 2001. Problem solving by heterogeneous agents. *Journal of Economic Theory*, 97, 123-163.
- Hong, L. and S. Page. 2004. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *PNAS*, 101:46, 16385–16389.

Kuorikoski, J. and Pöyhönen, S. (unpublished). Simulations as virtual experiments: an inferentialist approach. In review.

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

Longino, H. (1990). *Science as Social Knowledge*. Princeton: Princeton University Press.

March, J. 1991. Exploration and exploitation in organizational learning. *Organization Science*, 2:1, 71-87.

Mason, W. and D. Watts 2012. Collaborative learning in networks. *PNAS*, 109:3, 764–769

Mayo-Wilson, C., Zollman, K. J. S., & Danks, D. 2011. The independence thesis: When individual and social epistemology diverge. *Philosophy of Science*, 78:4, 653–677.

Merton, R. K. 1973. *The Sociology of Science. Theoretical and Empirical Investigations*. Chicago: The University of Chicago Press.

Muldoon, R. 2013. Diversity and the division of cognitive labor. *Philosophy Compass*, 8, 117-125.

Newell. A. & Simon, H. 1972. *Human Problem Solving*. Prentice Hall.

Piimies, J-P, and J. Sappinen (mimeo). The Advancement of science by means of economics. Unpublished manuscript, 2000.

Pöyhönen, S. 2016. Value of cognitive diversity in science. *Synthese*. Online first.

Riegler, A., & Douven, I. 2009. Extending the Hegselmann–Krause model III: From single beliefs to complex belief states. *Episteme*, 6:2, 145-163.

Rosenstock, S;C. O'Connor, and J. Bruner. Forthcoming. In epistemic communities, is less really more? *Philosophy of Science*.

Russell, S. J., and P. Norvig. 2003. *Artificial intelligence: A modern approach (2nd ed.)*. Prentice Hall series in artificial intelligence. Upper Saddle River, London: Prentice Hall/Pearson Education.

Strevens, M. 2003. The role of the priority rule in science. *The Journal of Philosophy*, 100:2, 55-79.

Sunstein, C. 2006. Deliberating groups versus prediction markets (or Hayek's challenge to Habermas). *Episteme*, 192-213.

Surowiecki, J. 2005. *The Wisdom of the Crowds*. London: Abacus.

Thoma, J. 2015. The Epistemic Division of Labor Revisited. *Philosophy of Science*, 82:3,454–472

Weisberg, M. 2010. New Approaches to the Division of Cognitive Labor, in P. D. Magnus & J. Busch (ed.), *New Waves in Philosophy of Science*. Palgrave Macmillan.

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

Zollman, K. J. S. 2010. The epistemic benefit of transient diversity. *Erkenntnis*, 72:1, 17–35.

Zollman, K. 2013. Network epistemology: Communication in epistemic communities. *Philosophy Compass*, 8, 15-27.