This version of the article has been accepted for publication, after peer review but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: http://dx.doi.org/10.1007/s13194-024-00601-9

Highly idealized models of scientific inquiry as conceptual systems

Author information

Renne Pesonen

renne.pesonen@gmail.com

ORCID: 0000-0001-6425-5772

Tampere University, the Faculty of Social Sciences

Kalevantie 4, 33014 Tampere, Finland

The social epistemology of science has adopted agent-based computer simulations as one of its core methods for investigating the dynamics of scientific inquiry. The epistemic status of these highly idealized models is currently under active debate in which they are often associated either with predictive or the argumentative functions. These two functions roughly correspond to interpreting simulations as virtual experiments or formalized thought experiments, respectively. This paper advances the argumentative account of modeling by proposing that models serve as a means to (re)conceptualize the macro-level dynamics of complex social epistemic interactions. I apply results from the epistemology of scientific modeling and the psychology of mental simulation to the ongoing debate in the social epistemology of science. Instead of considering simulation models as predictive devices, I view them as artifacts that exemplify abstract hypothetical properties of complex social epistemic processes in order to advance scientific understanding, hypothesis formation, and communication. Models need not be accurate representations to serve these purposes. They should be regarded as pragmatic cognitive tools that engender rather than replace intuitions in philosophical reasoning and argumentation. Furthermore, I aim to explain why the community tends to converge around few model templates: Since models have the potential to transform our intuitive comprehension of the subject of inquiry, successful models may literally capture the imagination of the modeling community.

1. Introduction

The high level of idealization in agent-based models of scientific inquiry has raised concerns about the epistemic function and utility of the methodology. These concerns become especially pressing if we treat simulations as virtual experiments (e.g., Fazelpour & Steel, 2022, 214), or use them for other purposes that require high-fidelity representation or prediction. However, scientific models have a wide range of uses that do not always require accurate representation and can even benefit from idealization, such as advancing scientific understanding, hypothesis formation, and communication. This paper argues that the key function of current agent-based models of scientific inquiry is to facilitate these three aims by supporting the development of cognitive skills and conceptual schemata. This claim is consistent with related suggestions that the proper use of simulation models of scientific inquiry is for theoretical exploration (Šešelja, 2021) or argumentation (Aydinonat, 2021). However, this paper does not prescribe how models should be used or evaluated. While my argument is prescriptive, it still aims to further answer the question of what the epistemic benefits of these highly idealized models are, if not prediction and accurate representation.

The methodological debate on modeling in social epistemology of science has largely focused on using formal models to investigate how to best design and modify the social organization of science. However, normative recommendations typically involve prediction, yet, the current models are neither calibrated nor validated against data (Martini & Pinto, 2017), their representational accuracy has been called into question (Thicke, 2020), and some have found them conceptually vague (Thompson, 2014; Bedessem, 2019). Some recent reviews have emphasized that simulation models of scientific inquiry are theoretical arguments rather than virtual experiments (Reijula & Kuorikoski, 2020; Aydinonat et al., 2021). Rather than for prediction, these models should be used, for example, to identify possible causal mechanisms and potential explanations (Šešelja, 2022). Furthermore, their evaluation necessitates assessing not only the relation between the model and its target but the wider argumentative context in which the model is used (Aydinonat et al., 2021).

In line with this argumentative approach, I focus on the pragmatic model-modeler and model-audience relationships, instead of the model-target relation. Rather than discussing how models of scientific inquiry should be used, my examination centers on how modeling influences the theoretical intuitions of modelers, and how it affects the research practices of the community engaged in model-based reasoning. Here, the notion of "theoretical" should be understood in a psychological sense as pertaining to a class of conceptions derived from inferential and communicative practices, while "intuition" refers to a fluent skill in using the acquired concepts for those practices. Following De Regt & Dieks (2005), I hold that a theory is intelligible to scientists if they can recognize its characteristic consequences without performing exact calculations.

The epistemological value of simulation models is often considered to lie in the possibility of explicating our conceptions about the target system (perhaps in an altered form), and then replacing our intuitive reasoning with the aid of the computer (e.g., Aydinonat et al., 2021; Mayo-Wilson & Zollman, 2021). I focus on the epistemological relation that basically runs in reverse. Instead of discussing how we externalize reasoning with computer models, I argue that we also internalize aspects of models and modeling practices through engaging in model-based reasoning. Therefore, the relationship between computer and mental models is best understood as interactive and complementary.

Rather than acknowledging the obvious fact that we use models to develop theories and arguments, I pursue a more nuanced hypothesis on how modeling can change our implicit understanding of the object of study, and consequently, the associated research practices and norms. There is a rich, psychologically informed literature on thought experiments and model-based reasoning in science. I find this literature useful for shedding light on this aspect of modeling within the social epistemology of science as well, where similar discussions have been largely absent. In particular, it has been often argued that models of scientific inquiry serve heuristic purposes, but the analysis of that function has thus far remained thin. Previous literature has also noted the heuristic benefits of recursive interactions among theory, model, and data. However, models of scientific inquiry are theoretical models that rarely connect with actual data. The mutually reinforcing loops between theory and models may not be solely virtuous, particularly if the theoretical nature and cognitive functions of models are not clearly understood while they begin to drive theoretical research.

Section 2 contains an overview of epistemic networks (Zollman, 2007) and epistemic landscape models (Weisberg & Muldoon, 2009), which are highly influential and illustrative examples of agent-based models of scientific inquiry. The three main topics—scientific understanding, hypothesis formation, and communication—are covered in Sections 3, 4, and 5, respectively.

Section 3 argues that pragmatic cognitive activity with model systems functions as a source of cognitive skills and intuitive conceptual schemata relevant to scientific reasoning and understanding. This section discusses the topic in general terms, not specifically focusing on modeling in social epistemology. Section 4 then discusses how the cognitive competencies acquired through experience with formal models guide hypothesis and theory formation in the context of current agent-based modeling in the social epistemology of science. The focus is on how the abstract properties of model systems guide the implicit comprehension and hypothesis formation concerning the opaque macro-dynamics of social epistemic interactions. This implicit guidance is essential for conceiving of models not only as explicit representational devices but also as sources of intuitive conceptual systems. Section 5 then explains how these hypotheses and theoretical conceptualizations, couched in terms of simulation models, affect and potentially benefit the wider research community beyond the actual modelers. Successful models recruit new users, who are not limited to modelers but include anyone who uses these models and modeling results for scientific reasoning. Idealization likely contributes to the success of models in this regard, since the gist of highly idealized models is easier to grasp, communicate, and adapt for new purposes compared to intricate simulations.

Thus, the plan of the paper is as follows: models equip modelers not only with external computational tools but also with mental models that guide scientific comprehension (Section 3). After formal models are utilized for theoretical argumentation concerning the dynamics of scientific inquiry (Section 4), the underlying mental models and conceptual intuitions they convey can be internalized by a wider research community through various cognitively efficient representational means (Section 5). Finally, Section 6 concludes.

2. Examples of agent-based models of scientific inquiry

My analysis focuses on agent-based models that explore the effects of diversity in scientific inquiry. Specifically, I discuss two well-known studies that have spurred a series of follow-up models and methodological debates on the use of simulation models in the philosophy of science. These models are the epistemic networks of Zollman (2007; 2010) and the epistemic landscape model of Weisberg and Muldoon (2009). The reason to focus on diversity models is that most simulation models of scientific inquiry fall into this category, or they are derived from models that investigate the effects of diversity in opinions, heuristics, or research strategies. Hence, this class of models is highly representative.

The epistemic networks (Zollman 2007) consist of agents who can investigate two competing theories. In each turn of the simulation, the agents can choose to sample either of the two options, A or B. In return, they receive either null or positive rewards. This cycle is conceptualized as conducting an experiment on either of the competing theories and receiving either a negative or positive result, such as a publishable experimental finding. The results are sampled from random and unknown distributions for each of the options, A and B. The agents' goal is to maximize their rewards in the long run. This means settling on the option that yields more positive outcomes, which is the correct theory or at least the one that is objectively more probable.

The agents are modeled as Bayesian learners who update their beliefs about the two theories as evidence accumulates. The agents face an elementary exploration/exploitation dilemma well-known in fields such as economics and artificial intelligence. The dilemma is that you want to exploit the option that has consistently produced rewards, but it is impossible to determine which option is superior without occasionally exploring the other one. In epistemic networks, the agents are not alone; they also monitor the actions and results of other agents. In Zollman’s (2007) seminal model, agents explore solely based on information received from others. They exploit their preferred option and switch only if the rival option gains more support within the network of agents they are monitoring.

Zollman demonstrated that the entire community can become locked into an inferior theory if promising but misleading results prematurely convince everyone. However, this problem can be mitigated by limiting the connectivity of the network. If agents monitor only few peers, the early results do not reach everyone. A part of the network may flip, but some agents may still persist with the other option and eventually persuade the rest. The diversity of beliefs, therefore, ensures that the community keeps viable alternatives available, at least transiently, even if they are not initially the best-supported by the evidence (Zollman, 2010).

The epistemic landscape (Weisberg & Muldoon, 2009) is an abstract representation of a research problem and its search-space. This landscape consists of a xy-plane of two discrete-valued dimension. These dimensions represent alternative “approaches” to the problem investigated, which may include alternative theories, experimental paradigms, or something alike. Each point on this xy-plane corresponds to a combination of approaches selected from each of the two dimensions. Each point is also associated with a measure of the epistemic significance of that specific combination. Some combinations are more fruitful than others, and agents navigate the landscape in an attempt to find the most significant combination

In the original Weisberg–Muldoon model, the landscape consists of a grid with two hills, and at the start of each simulation run, the agents are randomly positioned on it. The agents do not see the form of the landscape, but they rely on the combination of random walk and hill-climbing search heuristics. They move randomly one patch at a time. However, if they move on a patch of a higher significance than the one they had just left, they continue in the same direction. This process corresponds to a well-known learning scheme called gradient ascent, which involves a local search for incrementally better solutions. The challenge is that the agents tend to become trapped at local maxima instead of finding the global maximum, if they start their search on the slopes of the wrong hill. They can also get lost on a plain terrain that cannot guide the search.

The purpose of the epistemic landscape model is to investigate how social interaction influences the global search of the terrain when multiple agents are involved. Specifically, the model simulates the effects of diverse search strategies by incorporating three sub-populations: followers, mavericks, and controls. Each agent leaves a trace on the landscape. When follower agents encounter a trace, they follow it if it leads to higher significance in their local environment. Mavericks instead avoid previously visited patches and move away from trails left by others. Controls simply ignore the traces. Based on their simulations, Weisberg and Muldoon (2009) concluded that even a small number of mavericks significantly enhance the success of the entire population, as mavericks have the ability to discover hidden peaks of significance and guide others toward them.

My reason for focusing on these two studies is that they are widely known and representative examples of the use of agent-based models in the philosophy of science. They are particularly noteworthy for introducing novel ways to conceptualize the social dynamics of science, and they have inspired a subsequent generation of models that build upon their original frameworks. Some of the follow-up studies have modified the original model to examine the robustness of the initial results, while others have used the model templates to investigate new mechanisms and questions different from the original studies (see Aydinonat et al., 2021; Šešelja, 2022). Aydinonat et al. (2021) have tracked the evolution of epistemic landscape models, and argued that the modifications made to the original model and its versions can be interpreted as argumentative moves. I am more interested in what remains intact in these modifications, which is the abstract model template that serves as a conceptual foundation for a family of models used for various argumentative purposes.

Despite the differences of these two models, they share similarities. They both model well-known learning issues, namely the exploration/exploitation dilemma in epistemic networks and gradient ascent in epistemic landscapes. In addition, they both employ model templates with a history in other disciplines, implying that some conceptual and formal features of these models not only travel across research questions within social epistemology but also across disciplines. These features are addressed in Section 4 in more detail.

Before delving further into modeling in the social epistemology of science, the next section discusses mental modeling, scientific reasoning, and the psychology of comprehending complex causal scenarios. I argue that explicit computational models should not be understood merely as replacements for implicit mental models. Instead, their relationship is interactive; theoretical intuitions and cognitive skills develop through the coupling of internal and external models within cognitive activities that involve model systems.

3. Mental simulation and reasoning with models and thought experiments

In this section, I argue that while explicit simulation models help articulate theoretical intuitions, they also transform modelers' theoretical intuitions by providing new mental models, cognitive skills, and associated conceptual schemata. Moreover, explicit models help make abstract ideas accessible to mental imagery and other cognitive faculties capable of exploiting non-propositional knowledge. Therefore, model systems enable the development of new kinds of mental models and simulative reasoning. To explain how external and mental models are coupled, it is instructive to begin with a brief discussion of scientific thought experiments, as there is a rich body of epistemological and psychological literature exploring the connections between mental simulation, model-based reasoning, and thought experiments in science and philosophy.

I first provide a brief overview of the theory of mental models and mental simulation. Then, I explain how mental simulation is an integral part of cognitive skills, and how expertise facilitates the development of increasingly abstract conceptual schemata. These schemata enable intuitive understanding of the domain and potentially other domains via analogical transfer of knowledge. Finally, this section argues that model systems function as learning targets that allow the development of conceptual competencies relevant for scientific understanding. This claim is further utilized in Section 4, which argues that the cognitive influence of formal models on theoretical intuitions guides hypothesis formation in the social epistemology of science.

In the context of social epistemology, the comparison of computer simulations with thought experiments is sometimes made in passing (e.g., Currie & Avin, 2019; Aydinonat et al., 2021). Recently, Mayo-Wilson and Zollman (2021) provided a more thorough analysis of the idea, as they argued that philosophers should use simulations in place of thought experiments when reasoning about complex social interactions. This is because simulations can check our ideas about the dynamics of complex systems, which are often impossible to grasp intuitively. According to them, imagining a real-world process can also be understood as running a model. However, this is a mental model that remains implicit and untested, whereas computer simulations are explicit and can reveal unintuitive behaviors. Throughout this paper, I will rely on the idea that imagining a real-world process involves simulating a mental model of some kind.

Miščević (1992) and Nersessian (1992) were perhaps the first philosophers of science to propose that thought experiments are based on mental modeling. The exact nature of mental models is not entirely clear, but it is commonly held that they comprise mixtures of (quasi)perceptual mental imagery with causal and other forms of knowledge that preserve spatial and temporal features of the events from which the knowledge is derived. Mental simulation is a general cognitive mechanism intrinsic to understanding and coping with real-world phenomena by simulating scenarios represented by mental models. Much of the knowledge encoded in these models remains implicit and non-propositional, but thought experiments help make it accessible. Thought experiments exploit this simulative comprehension process by activating implicit knowledge through imagined scenarios, similar to how perception activates knowledge about how perceived events unfold.

Similarly, Gendler (2004) has argued that thought experimenting involves reasoning about a particular scenario, with a purpose to confirm or disconfirm a specific hypothesis or theory. In scientific contexts, the hypothesis concerns the features of the physical world, but the access to the scenario is via imagination instead of observation. According to Gendler, reasoning with mental images relies on non-propositional knowledge via quasi-perceptual processing, which is a fallible but reliable cognitive mechanism that we routinely rely on in everyday reasoning. With scientific thought experiments, we exploit this ordinary mental faculty with non-ordinary imagined scenarios that can reveal novel possibilities, inconsistencies, or other insights into some of our beliefs concerning real-world phenomena.

All three of these theorists hold that the cognitive mechanisms involved in the comprehension of and inference from thought experiments are ordinary cognitive faculties that we rely on in everyday reasoning. In particular, Nersessian's (2018) account of model-based mental simulation is rooted in a rich set of psychological mechanisms that exploit temporal and spatial structure, causal and procedural knowledge, mental animation, and more. Mental simulation is fundamentally a pragmatic faculty that has likely evolved to help us plan how to navigate the physical environment (Nersessian, 2018) and to anticipate the results of alternative courses of action, thus enabling the basic capacity for hypothetical reasoning (Koechlin, 2014). Apart from mental imagery, any general knowledge—from logic to folk and scientific theories—can be used to manipulate the simulated model and make further inferences based on it. Moreover, mental models and simulations can operate on multiple levels of specificity and abstraction (Barsalou, 1999; Nersessian, 2018).

The benefits of mental simulation are thought reside in the use of non-propositional knowledge that is often hard to verbalize and factor into explicit deductive reasoning. Moreover, some reasoning tasks are easier to carry out by using, e.g., spatial and visual models instead of propositional information. Imagistic processes may also be heuristically valuable as they enable novel and creative ways to combine and assess existing pieces of knowledge. In addition to these commonly identified features, the key benefit of simulative reasoning lies in its role in cognitive skills that allow us to surpass some cognitive limitations that hinder our ability to comprehend and reason with complex scenarios.

Intuitive inference and sound judgment develop through the learning of domain-specific cues, which allow familiar contexts to be simulated mentally (Klein, 1998; Kahneman & Klein, 2009). When experts face a problem in their domain of expertise, they need not pay attention to all the details, but they they focus on few broad strategic decisions and simulate one option at a time. If the simulated action feels credible, it is executed. If not, the simulation is either manipulated or replaced by another option (Klein, 1998). This intuitive competence is attributed to perceptual and procedural learning, where familiar contexts cue decisions from long-term memory that have proven successful in the past. Experts learn to chunk complex cues and procedures, and they learn to recognize goal-relevant patterns (Gobet & Chassy, 2009). Thus, expertise builds up by learning predictive and explanatory models that compress experiences and background knowledge into highly organized, perceptually cued intuitive know-how that can be activated via mental imagery. Therefore, imagined scenarios are able to recruit non-propositional knowledge in a cognitively efficient way and "mobilize the resources of highly organized background knowledge" (Miščević, 1992, 225), like recognition processes do in expert decision making.

This know-how also pertains to tasks associated with higher cognition and abstract reasoning, such as chess or doing philosophy. For example, when philosophers contemplate a Gettier case, their knowledge about the familiar use of the word "knowledge" is activated. This knowledge is neither explicit nor analytic, nor is it *a priori*. It is tacit, procedural, and derived from experience, as are cognitive skills in general. This highly learned procedural knowledge delivers intuitive understanding of how the word "knowledge" is properly used in the imagined scenario that many find to clash with the classical definition of knowledge.

Furthermore, the comprehension of familiar tasks gradually shifts to an increasingly abstract level from mere perceptual cues. Experts become more competent in projecting their knowledge across structurally similar problems and also across domains (Chi et al., 1981; Reeves & Weisberg, 1994). Thus, as the understanding of the domain becomes increasingly intuitive, the domain ceases to be a target of conceptual learning and it becomes a conceptual system in itself, enabling experts to (re)conceptualize and reason about other domains that bear schematic similarity to the concepts in their domain of expertise.

In the context of model-based research, this account explains how the interplay of internal and external models is crucial for scientific reasoning and understanding. Computational models, in particular, are good candidates for supporting perceptual, causal, and procedural learning, leading to the development of new cognitive skills and concepts that enable the intuitive simulation of hypothetical real-world processes.

For example, Nersessian (2008) traced historical records on how Maxwell derived his theory of electromagnetism. Maxwell relied heavily on visual diagrams and sketches that he continuously modified as his theoretical work progressed. These representations served as external models of the known properties of the electromagnetic forces, and their continuous manipulations—often under the guidance of various analogies—led to conceptual shifts on the dynamics and ontology of electromagnetism and eventually to the mathematical expression of the phenomena. A similar phenomenon is observed when graduate students in scientific and technical fields solve physics problems. Students produce and manipulate diagrams and exploit various visual aids and analogies to represent the dynamics of the target. They sketch and gesture over their external representations as they communicate and reason about the system (Nersessian, 2008).

The point of the above remarks is not to identify theories with models. Rather, the point is that models are cognitively efficient means to make sense of theories. Creative problem-solving in science involves model-based reasoning, where external and internal models are coupled through various interactions, supported and constrained by the model's features as well as available analogies and domain knowledge (see also Knuuttila, 2011; 2021). As tangible and manipulable representational tools, models facilitate the development of cognitive skills and the associated mental models and simulative model-based reasoning. Through this learning process, external models become internalized as intuitive conceptual systems that can be utilized for conceptualizing and reasoning about hypothetical real-world processes.

In conclusion, models have two epistemic functions: (1) to represent the ontology and dynamics of the target phenomena (in some relevant respects) and (2) to support intuitive understanding and creative reasoning about the target. These functions are not perfectly aligned, as representation can sometimes be independent of the aim of understanding, which is only hindered by excessive accuracy (Potochnik, 2015; Kuorikoski & Ylikoski, 2015). In the next section, I examine the epistemological import of agent-based simulation practices specifically in social epistemology, focusing on the latter function (2).

Computer models are sometimes understood to replace intuitive theoretical reasoning with automated inference over a particular formal representation of the target (e.g., Beisbart, 2012). In this section, I have argued that models should also be viewed as cognitive and communicative tools that function as sources of cognitive skills and conceptual intuitions. Experience with models and modeling techniques enables one to conceptualize other systems by these frameworks and to intuitively grasp the qualitative consequences of doing so. Essentially, this is analogical reasoning, where an actual or hypothetical real-world system is conceived and mentally simulated in terms of a familiar formal system. The next section further argues that often the theoretical ideas explicated in formal models are partly derived from model formalisms themselves. In particular, many agent-based models of scientific inquiry are partially based on content-neutral model templates that guide the conceptualization and formalization of the basic elements of the models.

4. Abstract model templates as conceptual systems

Much of the methodological discussion on modeling in the social epistemology of science has focused on the representational realism of models (Martini & Pinto, 2017; Thicke, 2020). Modeling studies usually build on existing ones, and new models are often advertised as more accurate than their predecessors in terms of the representational model-target relationship. In this section, I argue that current models represent modelers' theoretical ideas, which largely stem from the properties of model systems themselves. More specifically, I focus on the contribution of content-neutral model templates that, as such, do not represent anything specific. However, these templates already contain abstract conceptualizations of the target system and guide how model elements are conceptualized and formalized. I argue that abstract properties of models largely guide the concept and hypothesis formation in model-based research of scientific inquiry.

This claim is a natural implication of the argument in Section 3: Cognitive skills acquired through activities involving models train modelers to intuitively conceive poorly understood systems in terms of familiar model systems. The suggestion is not a criticism of modeling practices in the social epistemology of science, but rather an epistemological assessment of the consequences of how modeling influences hypothesis formation. The main argument of this paper is that this is one of the key epistemological functions of formal models of scientific inquiry. Section 5 further argues that it also benefits the wider research community by providing shared mental models that coordinate research and intersubjective theoretical intuitions.

Note that even if models represent only theoretical ideas derived from model systems, replacing thought experiments and mental simulation with computer simulations can convey significant epistemic benefits. Mental simulations are rooted in implicit and possibly idiosyncratic knowledge and experiences. According to Gilbert and Wilson (2007, 1354), “mental simulation is the means by which the brain discovers what it already knows.” Computer simulations have a very different epistemic nature. They are public, explicit, and formally well-defined, and they can uncover properties that are either unintuitive or too complex to be reliably grasped by theoretical or commonsense intuitions (see Mayo-Wilson & Zollman, 2021).

In this section, my aim is not to argue against analyses that emphasize the merits of computation over intuitive inference. Instead, I seek to identify and highlight a distinct epistemic function of these models. I argue that once we look beyond the mere computational aspect of agent-based simulation practices in the social epistemology of science, it appears that (pre)theoretical intuitions are not simply formalized and replaced by the explicit model and computation. With complex and causally opaque phenomena, such as social epistemic systems, model construction does not typically begin with a careful formal representation of the target’s components, followed by discovery of its properties through computation (cf. Mayo-Wilson & Zollman, 2021, 3663). Instead, modeling may usually begin by conceiving the target in terms of some abstract principle or model system and envisioning how the target would behave under that construal. This process involves analogical reasoning and mental simulation during theoretical ideation, where the abstract model formalism provides a conceptualization of the hypothetical macro-dynamics of the target.

If that is the case, modeling has two interacting cognitive aspects: (1) Experience with models and related formal tools enables researchers to conceive of target systems by the properties of these formalisms, and (2) constructing the actual model yields a public artifact that allows the modelers to computationally verify whether their intuitions hold in the actual model and under what conditions. The results attained in step (2) can be used for reasoning about the target, and they also feed back into the pool of knowledge exploited in step (1), potentially fueling further theoretical ideation based on the models and their templates.

This suggestion aligns with the Kuhnian account of normal science, which posits that scientific reasoning is typically case-based or model-based (Nickles, 2003). It does not proceed by general rules or theories but rather by established techniques for representing and inferring about specific research problems. The selection of a technique for a particular problem is guided by the problem's similarity to those with a known solution or approach. Thus, successful modeling efforts tend to reinforce the further use of the same or similar approaches for related research questions. This emerging pattern shapes how the domain of inquiry is conceived. Section 3 discussed how the comprehension of familiar tasks becomes increasingly abstract and how experts become more competent at projecting their knowledge across structurally similar problems and across domains. Hence, the similarity that guides the selection of an approach can be recognized on multiple levels of abstraction. This includes the high level of abstraction and idealization that many scientific models exhibit, explaining the prevalent reuse of the same abstract model templates for very different purposes, even across disciplines (see Gerrits & Marks, 2015; Knuuttila & Loettgers, 2016).

Before going into more detail, it is instructive to analyze model-based arguments in the social epistemology of science into more distinct components. The general form of simulation-based arguments in the social epistemology of science can be outlined roughly as follows: Given a modeling template T and its parametrization P, conditions C bring about an effect E. Here, T is the generic modeling template or formal framework used (epistemic landscapes, epistemic networks, etc.), and parametrization P fixes the structural assumptions and variables of the model. Conditions C specify what variables are investigated in the actual simulation runs, and E is the interesting property discovered. How E depends on C is discussed in order to support the actual theoretical argument of the study, and some reasons are provided why the particular framework and parametrization was chosen. The formal dependencies in the simulation model constitute a chain $T \rightarrow P \rightarrow C \rightarrow E$, where each step is constrained by the preceding one and the last step is computationally determined. Step $T \rightarrow P$ corresponds to the construction of the formal model and $C \rightarrow E$ to the simulation runs and their interpretation. Note that the purpose of this characterization is not to depict the actual model construction process but solely to facilitate the discussion below.

For concreteness, here are two examples of the above argument form. First, If we use epistemic networks to represent the exploration and distribution of information among scientists (T), and we assume that scientists always use the hypothesis they believe has the best support and explore only through information received from peers (P), then restricting the connectivity in the networks (C) will prevent the community from prematurely locking into an inferior hypothesis (E) (Zollman 2010). Second, if we use fitness landscapes to represent the epistemic success of research as a function of combinations of different approaches (T), and we assume that progress of individual scientist can be modeled by gradient ascent that is sensitive to the already visited locations of the landscape (*P*), then increasing the proportion of agents who avoid already covered areas (*C*) will increase the epistemic success of the community (*E*) (Weisberg & Muldoon, 2009).

The findings concerning the relation $C \rightarrow E$ are usually described as the main contribution of the model. This is also where the added epistemic value of the simulation runs is considered to reside, i.e., in the uncovering of the effects of model variables through numerical computation, especially when complexity trumps analytical and intuitive inferences. Robustness analyses also target the relation $C \rightarrow E$ by extending the simulated parameter space in *C. S*ome robustness analyses also focus on P, involving the addition, removal, or modification of mechanisms and idealizations in the simulation (see Kuorikoski & Ylikoski, 2015; Frey & Šešelja, 2018). Changes in the parametrization are usually interpreted as modifications to the existing model, while changing the underlying template means to replace the entire model with a qualitatively different kind.

What interests me here is not the computational process from *C* to *E*, but rather the conceptualizing process with *T* and *P*, which together give the formal structure and interpretation of the model. Following the analysis by Aydinonat et al. (2021), the model template is involved in setting up a *conceptual model,* which lays down the qualitative assumptions about the modeled phenomenon and which guides the parametrization and construction of the actual computational model. The authors did not fully specify the role of the template. However, they argued that the conceptual model conveys a set of hypothetical properties of the modeled target. These are made precise, albeit often in idealized or distorted form, in the parametrization.

However, the template does not usually function as a rough representation of the modeled real-world target, which is subsequently made more precise by the parameters. In scientific modeling, it is common for the template to import some generic abstract property, or a set of properties that are independent of any subject-specific interpretation of the model. The interpretation is provided alongside the parameters chosen for that particular modeling purpose (Knuuttila, 2011; Knuuttila & Loettgers, 2016). Thus, the template already conveys an abstract idea of the hypothetical structure or dynamics of the target. Hence, if the template is involved in devising of the conceptual model, the abstract properties associated with the template are an integral part of the initial conceptualization of the target. In that case, modeling proceeds in a top-down manner, starting from the abstract conceptualization of the dynamics and moving on to the specification of more detailed elements, such as agents and their actions in agent-based models.

In current simulation studies in social epistemology, the justifications for choosing specific templates are often quite elliptic. For example, it is unclear what the epistemic landscape and its dimensions actually represent (Bedessem, 2019). Similar criticisms have been directed toward, e.g., Hong and Page's (2004) diversity model (Thompson, 2014). It is also unclear why Hong and Page chose that particular model for investigating their famous diversity-trumps-ability result instead of using their earlier template, which, in fact, does not support the result (Reijula & Kuorikoski, 2021). Some models are more semantically transparent. For instance, it is arguably easier to understand how the elements in Zollman's (2007; 2010) epistemic networks map to scientific experimentation and communication. However, even these models are rough idealizations that omit many features prominent in theory choice for scientists, such as argumentation and the role of funding allocations for different research programs.

However, in scientific modeling, it is common and often necessary to focus on factors beyond accurate representation. Specific templates may be chosen not only for their representational adequacy but also for pragmatic reasons, including tractability and familiarity. These qualities make the properties of the template salient and heuristically useful to the modeler during hypothesis formation and model design. Familiarity may arise from domain expertise with modeling or from the template's earlier successful adaptations in various fields (Knuuttila & Loettgers, 2016; see also Gerrits & Marks, 2015).

I suggest that in the current models used in the social epistemology of science, the choices of modeling frameworks are based more on their heuristic role and salience than on representational accuracy. This implies that when modelers specify and justify their modeling assumptions, their theoretical intuitions may reflect more the abstract properties of model systems than the presumed empirical properties of their real-world targets. If a template has a history, it can be selected based on its known behavior, i.e., the qualitative understanding of the relation $C \rightarrow E$, if it appears conceptually relevant for the current modeling task. The model template is then adapted for the current modeling task by giving it an interpretation that guides its parametrization, and the computational runs verify which interesting properties survive the adaptation and which new ones emerge (see Knuuttila, 2011, 268). Some of the key findings may then be more properly considered as being replicated in the adaptation rather than discovered through computation, and the theoretical discovery resides largely in the adaptation process wherein the intuitive understanding about the model's behavior meshes with (pre)theoretical conceptions of the target in creative ways.

For example, epistemic landscapes can be traced to fitness landscape models in biology, from where they have been adapted for a wide range of purposes in disciplines including sociology, economics, law, and anthropology (Gerrits & Marks, 2015). While Weisberg & Muldoon (2009) did not mention the history of the model template, in some subsequent works this connection is made explicit (e.g., Alexander et al., 2015).

Zollman has been more explicit on the origins of his epistemic network models as well as theoretical and analogical reasoning behind them. Epistemic networks were adapted from a model of social learning of economic choice (Bala & Goyal, 1998). The pioneering paper that introduced these networks to model scientific interaction highlighted the similarities between the economics model and scientific inquiry (Zollman, 2007). In the original Bala–Goyal model, agents face an uncertain recurring choice, and they are sensitive to the information received from others and upgrade their behavior accordingly. This resembles how scientist respond to socially transmitted evidence within their professional networks. In a subsequent publication, Zollman (2010) further discussed how his modeling results bear similarities to well-known theoretical ideas about the benefits of the diversity of opinions in science.

Epistemic networks in particular exemplify the general point of this section. Zollman's networks are based on an abstract but intuitive analogy between scientists and the populations simulated by Bala and Goyal. Based on this analogy, the existing formal model guided the selection of the relational and functional properties to be imported from the socio-economic study – even if the resulting model omitted important features specific to scientific inquiry. The first published study (Zollman, 2007) qualitatively replicated the simulation results by Bala and Goyal. These results, despite the omissions, proved to be analogous with influential theories in the philosophy of science and, hence, conceptually highly relevant.

Some agent-based simulations in social epistemology appear to be based on novel templates, however. For example, Hong and Page (2001; 2004) have introduced two such models. Hegselmann and Krause's (2009) model of opinion dynamics is also an example. The early publications that explain the motivation and construction of these models (Hong & Page, 1998; 2001; Hegselmann & Krause, 2002) rely far more heavily on modeling literature than detailed empirical considerations of the mechanisms and variables of the social phenomena being simulated. The theoretical discussion about the empirical subject matter proceeds on a very general and intuitive level as the authors explain the rationales for developing their modeling frameworks.

Therefore, the construction of these simulations was apparently guided by the experience with formal systems more than careful empirical considerations of their purported real-world targets. To be clear, I am not suggesting that there is something wrong with this. My point is only that even when the current models of scientific inquiry are not adapted from existing templates, their construction seem to be guided by abstract hypothetical conceptualizations, and they are not even intended to be accurate representations of any real-word systems. Hegselmann and Krause (2009, 131) are particularly explicit in this. Instead of being descriptions of actual events and processes, simulations in social epistemology are means to investigate the robustness of the general logic of arguments (see Hong and Page, 1998, 17).

In this section, I have relied on accounts of scientific modeling that share the consensus that models are generally not accurate representations of the ontology and laws of their targets. Instead, they are cognitive tools that support reasoning about selected aspects of modeled phenomena by instantiating abstract formal systems for surrogate reasoning (Knuuttila, 2011; 2021; Kuorikoski & Ylikoski, 2015; Currie, 2019; Aydinonat et al., 2021). Scientific models, in general, have various purposes, some of which require accurate representation. However, the account adopted here seems particularly adequate for describing the current use of agent-based models in the social epistemology of science. These highly idealized models conceptualize the hypothetical macro-dynamics of social epistemic interactions in ways that are largely derived from abstract properties of model formalisms themselves. Because they are theoretical models, modeling results reveal the consequences of these conceptualizations, not real-world events.

These conclusions are not offered to reveal shortcomings of current models of scientific inquiry, but rather their proper epistemic interpretation. The ability to conceptualize the opaque dynamics of complex social interactions in cognitively efficient ways is a major cognitive benefit not only for the modeler but for the wider research community. The next section discusses this aspect not solely in terms of modeling but in terms of model-based argumentative and theoretical reasoning more generally.

5. Collaborative reasoning with intuitive models

In this section, I argue that the cognitive impact of model-based reasoning is not exclusive to actual modelers. Section 3 argued that model systems, as tangible external tools, enable pragmatic interactions that facilitate the development of new cognitive skills and conceptual schemata relevant for scientific reasoning and understanding. More specifically, this understanding is rooted in the mental simulation of cognitive activities that involve model systems. While mental simulation is associated with quasi-perceptual mental imagery, what is more crucial is the experience accumulated through pragmatic interaction, which fosters the development of new cognitive skills and conceptual schemata that can be recruited to reason about other domains via the analogical transfer of skills and knowledge. In Section 4, I argued that these skills guide modelers to conceptualize and reason about real-world phenomena in terms of familiar model systems and their abstract properties. Here, I argue that engaging in actual modeling is not necessary for this conceptualization function to have its cognitive impact. It suffices for researchers to utilize models and their properties as conceptual systems for scientific reasoning. While allowing highly idealized models to drive argumentative practices involves cognitive risks, they also offer cognitive benefits by providing shared mental models that coordinate collaborative research.

It is instructive to note the broad similarities between the heuristic use of models and the psychology of analogical reasoning: An unfamiliar target problem or scenario is compared against a set of potential analogies, and some form of similarity is recognized between the target and a candidate source analogy. Irrelevant surface properties are excluded from consideration, and the underlying “deep structure” of the source, often a causal or other relational structure, is mapped onto the target and adapted for a better fit for its constraints. The acquired scenario is run (mentally or on a computer), and the outcome is interpreted to answer our questions about the target. Essentially, the same general idea characterizes both modeling in the social epistemology of science and analogical reasoning with mental models (see Holyoak et al., 2010; Reeves & Weisberg, 1994). In the analogical reasoning literature, it is commonly held that the source analogy is selected based on its similarity to the deep structure of the target. However, if we already knew the target structure, there would be no need to import it from an analogy (Reeves & Weisberg, 1994). Analogies are useful precisely when we do not understand how the target functions, and hence we need to conceptualize it with the aid of familiar schemas or scenarios. This is the main epistemic function of highly idealized models of scientific inquiry.

In the previous two Sections, I argued that modelers are able to exploit the abstract properties of model systems for conceptualizing hypothetical real-world processes. However, there is no reason to suppose that the ability to use models for surrogate reasoning is within the reach of only modelers. The skilled use of models involves understanding how they function, which is derived from toying and experimenting with them. However, it also involves the ability to use them for inferential and communicative practices concerning actual and hypothetical systems. While modelers are best positioned to acquire the former kind of skills, engaging in modeling is not necessary to learn expertise in using models and modeling results for scientific reasoning and communication. These represent two related but distinct cognitive skill sets. Using models for scientific reasoning requires learning the relevant inferential norms and practices prevalent in the research community, which stem from conceptualizing hypothetical real-world phenomena with models.

The reason for emphasizing this is that modelers evidently acquire distinct cognitive skills and conceptual competencies. However, it is less clear whether this applies to other members of the research community. Perhaps they only absorb the modeling results in the form of explicit hypothetical claims without significant conceptual or procedural learning. Nonetheless, I believe this is not the case. It is by no means trivial to understand how the highly idealized models of scientific inquiry are supposed to map onto their purported real-world targets and what, if anything, can be learned from them. Understanding modeling practices and model-based arguments requires conceptual learning and the adjustment of theoretical intuitions to align with the properties of model systems and aims of modeling. Some of the non-propositional knowledge generated by modeling can be shared across the research community through diagrams and other means. These methods convey the model's structure and dynamics in heuristically useful and cognitively efficient ways. Consequently, models can provide different skills and concepts to different researchers, depending on how they learn to reason with models and the representational means they use to access the modeling results.

This difference in competence between modelers and the rest of the community can be understood along the lines of Harry Collins’ (2004) distinction between contributory and interactive expertise. Contributory expertise refers to the ability to exercise a practical skill, while interactive expertise refers to the ability to expertly converse about the skill without being able to exercise it. In this context, these forms of expertise can be seen as the skill to program and use computational models and the skill to competently discuss and reason about them. However, competent utilization scientific models involves not only reasoning about models but also with models in order to devise and understand model-based arguments. Anyone who is proficient in using models and modeling results for hypothesis formation, inference, and argumentation, possesses contributory expertise to that extent.

Apart from these psychological aspects, there is also a social coordinative factor that is crucial for the conceptualization function of models. In Section 4, I suggested that successful modeling efforts contribute back to the pool of shared knowledge that enables the ideation of theoretical models in the first place. The significance of this factor becomes most evident when the initial model serves as a foundation that coordinates further theoretical work through broader model-based reasoning in addition to actual model construction and simulation.

Computational models of scientific inquiry are introduced for investigating specific effects. Subsequent discussions around the seminal results mostly retain the basic template, and the argumentative moves in the following debates revolve around adding and tweaking the model parameters. These investigations into new or corrected mechanisms are common in the ensuing critical assessment of models. Critiques of specific models are plentiful (Thicke, 2020), and not all critiques assess only the representational details but deeper conceptual issues involved in the model templates (e.g., Thompson, 2014; Bedessem, 2019). However, rarely are the entire templates replaced while keeping the original research questions intact. This would amount to a more radical reconceptualization of the phenomenon under study, but subsequent research tends to retain and build upon the original formal conceptualization. Often, this reuse of templates also occurs when research questions change. Consequently, successful templates sometimes evolve into more general tools for representing scientific inquiry, as they are adapted for diverse purposes beyond the scope of the original research questions (see Aydinonat et al., 2021; Šešelja, 2022).

As argued in Section 4, I suggest that the reason for sticking with successful templates, and allowing them to drive the research, is not solely it is easier than developing a new one for each modeling purpose. Scientific reasoning in general is based on the selection and adaptation of familiar techniques within one’s domain of expertise. Consequently, the use of particular frameworks tends to reinforce their further use for related questions. In the context of social epistemology, this tendency trains the participants to think of scientific inquiry in terms of the familiar model systems and their properties. In doing so, they further entrench the ways in which the research community conceptualizes their research questions in terms of model system, shaping the argumentative practices and the norms concerning competent research.

In particular, if a model draws attention, critics draw upon their own understanding of the modeled phenomena to point out how the model misses some relevant mechanisms or formalizes them incorrectly, how it is based on unfounded assumptions, or how it could be better operationalized. Some may be less critical and utilize its resources for further theoretical arguments. This is ordinary scientific reasoning in the course of normal science, which now revolves around the conceptualization brought by the initial model.

One prominent example of the development of a theoretical idea through iterated modeling efforts is the concept of transient diversity of opinions. As explained in Section 2, the idea is that the research community benefits from investigating several theories, even if only one of them is best supported by evidence. This raises the chances that the community eventually settles on the correct theory. Zollman (2010) first proposed that transient diversity could be maintained by restricting the information flow among scientists or by biasing them with extreme prior opinions. Later, Kummerfeld and Zollman (2016) discovered that transient diversity prevails if individual scientists occasionally test different theories instead of only their favorite one. Fazelpour and Steel (2022) demonstrated that the same benefit can result from in-group/out-group biases among scientists. Frey and Šešelja (2018) also found that the transient diversity can be secured by inertia that delays the theory switch for individual scientists.

If one presumes that the purpose of modeling scientific inquiry is to identify unique mechanisms for specific macro-scale phenomena, this may seem like a conflicting collection of results, as it remains unclear which mechanism is the key factor in transient diversity. However, this is usually not the aim. While the results might reduce the credibility of any single mechanism as being the main factor for transient diversity, together they reveal two valuable insights: first, transient diversity is a multifaceted phenomenon that can be implemented by various mechanisms; second, these modeling results suggest that the benefit of transient diversity is a robust abstract principle, regardless of its implementation. Moreover, any singular finding might be an artifact of the specific model implementation, so the proper unit of analysis should not be a single model but rather a whole family of models investigating the same phenomenon (Kuorikoski & Ylikoski, 2015). While all the simulation results referenced above were based on modifications of the epistemic network template introduced by Zollman (2007), there is no obvious reason for these investigations to adhere to the same template. For example, Šešelja (2022) noted that an argumentative agent-based model also replicated the diversity result, and essentially for the same abstract reason as the Kummerfeld and Zollman's (2016) model, even though it was based on a very different template.

There is nothing wrong with reusing existing templates for theoretical exploration, even if they are highly idealizing. After all, all the assumptions are made explicit in the models. However, there may be a more subtle effect: models that drive research may implicitly guide researchers to imagine complex real-world phenomena in terms of these idealizing systems when conceiving how real-world systems actually behave. For instance, current models are frequently used for making normative recommendations (see Thompson, 2014; Martini & Pinto, 2017; Bedessem, 2019; Thicke, 2020). But suggestions to take action in the real world cannot be based solely on theoretical ideation; it must be assumed that the models inform us of the expected consequences of actual interventions. However, current models are mostly adequate only for theoretical exploration. While many modelers pursue empirical adequacy by increasing complexity, making the current models more complex does not necessarily make them more reliable representations of real-world events (Martini & Pinto, 2017; Šešelja, 2021).

Therefore, the tendency of model systems to capture the imagination of the research community may not be solely advantageous. However, it does offer a clearly identifiable epistemic benefit: The sources of our intuitions about the dynamics of complex social systems are likely to be highly diverse and dependent on personal backgrounds and theoretical orientations. For example, Marxist sociologists and evolutionary psychologists may have very different understandings of the dynamics and key variables of social interactions. Additionally, we all have experience-based intuitions about individual behavior, which we likely extrapolate to our mental models of large-scale social systems. One commonly identified benefit of even highly idealized models is that, as public and well-defined systems, they are capable of being accessed, evaluated, and refined by others. If our conceptual understanding is refined through the use of models, model systems also serve as a means to coordinate *shared* understanding within the research community regarding the opaque macro-dynamics of social epistemic processes. In this role, simulation models do not regiment our intuitions merely as computational systems but as shared conceptual systems. As argued in this section, this epistemic benefit is available for the whole research community and not solely for modelers.

6. Conclusions

I have expanded the argumentative account of modeling in the social epistemology of science by examining the cognitive functions of agent-based models as conceptual systems. Specifically, I have argued that the use of models facilitates the development of cognitive skills and conceptual schemata that allow modelers to mentally simulate real-world processes by exploiting their understanding of model systems and their abstract properties as sources of theoretical intuitions. Hence, computational models serve not only as explications or replacements for intuitive mental modeling but also provide material for it. I explored this aspect specifically within the social epistemology of science, extending the discussion to include the socially coordinative functions of model-based reasoning and its cognitive impact on users beyond the actual modelers.

My argument includes both positive and critical contributions. I have primarily focused on the positive contribution of explaining how model systems support intuitive reasoning and understanding, where intuition has a specific psychological interpretation as mental simulation based on mental models. I have argued that cognitive skill-learning and expertise provide the correct theoretical frameworks for understanding the cognitive impact and the social epistemology of modeling and model-based reasoning practices. While highly idealized models simplify and distort, idealization facilitates understanding and concept formation, and, consequently, communication, argumentation, and intuitive model-based reasoning.

The critical implications of my argument are mostly latent. The learning that facilitates the aforementioned cognitive benefits targets model systems, not real-world phenomena. Separate arguments are needed to justify the relevance of modeling results if their purpose is to make empirical claims. However, if models are used solely to demonstrate the qualitative consequences of a particular conceptualization of the target process, they should be evaluated similarly to analogies. The success of analogies is primarily a matter of pragmatics rather than veridicality or accurate representation. The source analogy serves its purpose if it promotes hypothesis formation, correct inferences, mutual understanding, and orderly communication and argumentation.

Computational modeling offers the clear epistemological benefit of forcing us to precisely articulate and formalize our every assumption (Mayo-Wilson & Zollman, 2021, 3663) Nevertheless, the necessary idealizations in the computational models of scientific inquiry also force us to simplify and misrepresent, leaving the mapping between the model and its supposed real-world target occasionally obscure. Even very remote analogies (or highly idealized models) can convey an abstract gist or structure that they actually share with the target. However, when this is not the case, both analogies and idealized models can be misleading and cause a false sense of understanding. In such instances, certain conceptualizations may pose high epistemological risks if the formal framework itself drives theoretical research on ill-understood phenomena.

This approach to the epistemology of simulation models may appear to downplay their computational aspect. However, the computational properties of models are essential for constituting the associated conceptual systems because they establish how the models actually behave. This, in turn, determines the characteristic consequences of conceptualizing the target through a particular model system. Computational runs also provide systematic feedback to modelers about the consequences of changes in model parameters. This factor is crucial since a predictable environment and feedback are prerequisites for the development of any complex intuitive expertise (Kahneman & Klein, 2009). Simulation models are unique cognitive tools precisely because they are interactive causal systems and not mere representations of causal systems.

To be clear, I do not deny that agent-based models of scientific inquiry can also boost our epistemic practices as computational tools. Neither do I deny that genuine discoveries through computation occur. I have simply analyzed modeling from a broader perspective, emphasizing aspects beyond the computational part. However, the reason why this part is particularly important in the social epistemology of science is that it is challenging to establish clear and agreed-upon conceptualizations of such a complex process as the social dynamics of scientific inquiry. Highly idealized models help overcome this problem by providing relatively accessible conceptualizations for facilitating scientific reasoning and communication. With regard to these aims as well as hypothesis formation, argumentation, and mutual understanding, idealization is a epistemic virtue rather than a vice, as far as it promotes comprehensibility.

The claim that highly idealized models function as conceptual systems is quite general and involves more specific functions such as explanation, argumentation, and the coordination of non-propositional knowledge. I have elaborated the claim by discussing the psychology of model-based and analogical reasoning, mental simulation, and cognitive skills. I hope that accounting for these cognitive factors helps us understand the epistemological prospects and potential problems of highly idealized theoretical models, especially in the context of modeling practices in the social epistemology of science.

References

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.

Aydinonat, N. E., Reijula, S., & Ylikoski, P. (2021). Argumentative landscapes: the function of models in social epistemology. *Synthese, 199*, 369–395.

Bala, V., & Goyal, S. (1998). Learning from Neighbors. *Review of Economic Studies 65*(3), 595–621.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(4), 577–660.

Bedessem, B. (2019). The division of cognitive labor: two missing dimensions of the debate. *European Journal for Philosophy of Science, 9*:3, https://doi.org/10.1007/s13194-018-0230-8.

Beisbart, C. (2012). How can computer simulations produce new knowledge? *European Journal for Philosophy of Science, 2*(3), 395–434.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices. *Cognitive Science, 5*(2), 121–152.

Collins, H. (2004). Interactional expertise as a third kind of knowledge. *Phenomenology and the Cognitive Sciences, 3*(2), 125–143.

Currie, A., & Avin, S. (2019). Method Pluralism, Method Mismatch, & Method Bias. *Philosophers' Imprint, 19*(13), 11–22.

Fazelpour, S., & Steel, D. (2022). Diversity, Trust, and Conformity: A Simulation Study. *Philosophy of Science, 89*(2), 209–231.

Frey, D., & Šešelja, D. (2018). What Is the Epistemic Function of Highly Idealized Agent-Based Models of Scientific Inquiry? *Philosophy of the Social Sciences, 48*(4), 407–433. https://doi.org/10.1177/0048393118767085

Gendler, T. S. (2004). Thought Experiments Rethought – and Reperceived. *Philosophy of Science, 71*(5), 1152–1163.

Gerrits, L., & Marks, P. (2015). The evolution of Wright's (1932) adaptive field to contemporary interpretations and uses of fitness landscapes in the social sciences. *Biology & Philosophy, 30*(4), 459–479.

Gilbert, D. T., & Wilson, T. D. (2007). Prospection: Experiencing the Future. *Science, 317*(5843), 1351–1354.

Gobet, F., & Chassy, P. (2009). Expertise and Intuition: A Tale of Three Theories. *Minds & Machines, 19*(2), 151–180.

Hegselmann, R., & Krause, U. (2002). Opinion Dynamics and Bounded Confidence Models, Analysis, and Simulation. *Journal of Artificial Societies and Social Simulation, 5*(3), https://www.jasss.org/5/3/2/2.pdf

Hegselmann, R., & Krause, U. (2009). Deliberative Exchange, Truth, and Cognitive Division of Labour: A Low-Resolution Modeling Approach. *Episteme, 6*(2), 130–144.

Holyoak, K. J., Lee, H. S., & Lu, Hongjing (2010). Analogical and Category-Based Inference: A Theoretical Integration With Bayesian Causal Models. *Journal of Experimental Psychology: General 139*(4), 702–727.

Hong, L., & Page, S. E. (1998). Diversity and Optimality. SFI working paper. Retrieved February, 16, 2023, from https://www.santafe.edu/research/results/working-papers/diversity-and-optimality

Hong, L., & Page, S. E. (2001). Problem Solving by Heterogeneous Agents. *Journal of Economic Theory, 97*(1), 123–163.

Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences, 101*(46), 16385–16389.

Kahneman, D., & Klein, G. (2009). Conditions for Intuitive Expertise: A Failure to Disagree. *American Psychologist 64*(6), 515–526.

Klein, G. A. (1998). *Sources of Power: How People Make Decisions*. The MIT Press.

Knuuttila, T. (2011). Modelling and representing: An artefactual approach to model-based representation. *Studies in History and Philosophy of Science, 42*(2), 262–271.

Knuuttila, T. (2021). Imagination extended and embedded: artifactual versus fictional accounts of models. *Synthese, 198*(Suppl 21), S5077–S5097.

Knuuttila, T., & Loettgers, A. (2016). Model templates within and between disciplines: from magnets to gases – and socio-economic systems. *European Journal for Philosophy of Science, 6*(3), 377–400.

Koechlin, E. (2014). An evolutionary computational theory of prefrontal executive function in decision-making. *Philosophical Transactions of the Royal Society B, 369*(1655). https://doi.org/10.1098/rstb.2013.0474.

Kummerfeld, E., & Zollman, K. J. S. (2016). Conservatism and the Scientific State of Nature. *The British Journal for the Philosophy of Science*, *67*(4), 1057–1076.

Kuorikoski, J. & Ylikoski, P. (2015). External representations and scientific understanding. *Synthese, 192*(12), 3817–3837.

Martini, C., & Pinto, M. F. (2017). Modeling the social organization of science: Chasing complexity through simulations. *European Journal for Philosophy of Science*, *7*(2), 221–238.

Mayo-Wilson, C.. & Zollman, K. J. S. (2021). The computational philosophy: simulations as a core philosophical method. *Synthese*, *199*(1–2), 3647–3673.

Miščević, N. (1992). Mental models and thought experiments. *International Studies in the Philosophy of Science*, *6*(3), 215–226.

Nersessian, N. J. (1992). In the Theoretician's Laboratory: Thought Experimenting as Mental Modeling. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, 1992*(2), 291–301.

Nersessian, N. J. (2008). *Creating Scientific Concepts*. The MIT Press.

Nersessian, N. J. (2018). Cognitive Science, Mental Modeling, and Thought Experiments. In: M. T. Stuart, Y. Fehige, & J. R. Brown (Eds.), *The Routledge Companion to Thought Experiments* (pp. 309–326). Routledge.

Nickles, T. (2003). Norman Science: From Logic to Case-Based and Model-Based Reasoning. In: T. Nickles (Ed.), *Thomas Kuhn* (pp. 142–177). Cambridge University Press.

Potochnik, A. (2015). The diverse aims of science. *Studies in History and Philosophy of Science Part A*, *53*, 71–80.

Reeves, L. M., & Weisberg, R. W. (1994). The Role of Content and Abstract Information in Analogical Transfer. *Psychological Bulletin, 115*(3), 381–400.

Reijula, S., & Kuorikoski, J. (2020). Modeling Epistemic Communities. In: M. Fricker, P. J. Graham, D. Henderson, & N. J. L. L. Pedersen (Eds.) *The Routledge Handbook of Social Epistemology* (pp. 240–249). Routledge.

Reijula, S., & Kuorikoski, J. (2021). The Diversity-Ability Trade-Off in Scientific Problem Solving. *Philosophy of Science 88*(5), 894–905.

Šešelja, D. (2021). Exploring Scientific Inquiry via Agent-Based Modelling. *Perspectives on Science, 29*(4), 537–557.

Šešelja, D. (2022). Agent-based models of scientific interaction. *Philosophy Compass, 17*(7):e12855. https://doi.org/10.1111/phc3.12855.

Thicke, M. (2020). Evaluating Formal Models of Science. *Journal for General Philosophy of Science, 51*(2), 315–335.

Thompson, A. (2014). Does Diversity Trump Ability? An Example of the Misuse of Mathematics in the Social Sciences. *Notices of the American Mathematical Society 61*(9), 1024–1030.

Weisberg, M., & Muldoon, R. (2009). Epistemic Landscapes and the Division of Cognitive Labor. *Philosophy of Science, 76*(2), 225–252.

Zollman, K. J. S. (2007). The Communication Structure of Epistemic Communities. *Philosophy of Science, 74*(5), 574–587.

Zollman, K. J. S. (2010). The Epistemic Benefit of Transient Diversity. *Erkenntnis, 72*(1), 17–35.