



Formal models of the scientific community and the value-ladenness of science

Vincenzo Politi¹

Received: 26 May 2020 / Accepted: 27 September 2021
© The Author(s) 2021

Abstract

In the past few years, social epistemologists have developed several formal models of the social organisation of science. While their robustness and representational adequacy has been analysed at length, the function of these models has begun to be discussed in more general terms only recently. In this article, I will interpret many of the current formal models of the scientific community as representing the latest development of what I will call the ‘Kuhnian project’. These models share with Kuhn a number of questions about the relation between individuals and communities. At the same time, they also inherit some of Kuhn’s problematic characterisations of the scientific community. In particular, current models of the social organisation of science represent the scientific community as essentially value-free. This may put into question both their representational adequacy and their normative ambitions. In the end, it will be shown that the discussion on the formal models of the scientific community may contribute in fruitful ways to the ongoing debates on value judgements in science.

Keywords Social epistemology of science · Agent-Based Models · Non-epistemic values in science · Value-free ideal · Thomas Kuhn

1 Introduction

In the past few years, social epistemologists have developed a number of formal models of the social organisation of scientific research—or, to borrow from Martini and Fernández Pinto (2017), ‘SOSR models’. Such models provide highly idealised representations of the scientific community. Overviews on the most discussed SOSR models in philosophy of science have been provided by, among the others, Weisberg (2010) and Reijula and Kuorikoski (2019).

✉ Vincenzo Politi
vin.politi@gmail.com; vincenzo.politi@medisin.uio.no

¹ Faculty of Medicine, Department of Biostatistics, Institute of Basic Medical Sciences, University of Oslo, IKlaus TorgErds vei 3, 0372 Oslo, Norway

The philosophical literature often focuses on technical issues with specific SOSR models. While their robustness and representational adequacy have already been analysed at length, however, there are many other questions about SOSR models in general which are yet to be answered. For instance, what is the ‘philosophical argument’ that these models are making? What can philosophers, scientists, and policy makers learn from SOSR modelling? Are computer based methods the best ways to represent the social organisation of science? This article is a contribution to the philosophical debate on SOSR models *in general*. Its aim is to understand which view of science these models end up supporting.

In this article, I will interpret many of the current SOSR models as being part of what I will call the ‘Kuhnian project’ in philosophy of science, which consists in the philosophical investigation of the scientific community. Apart from looking at the scientific community as the object of their inquiry, it will be shown that SOSR models share with Kuhn a number of questions about the relation between individual scientists and the collectivity they belong to. At the same time, however, they also inherit some of Kuhn’s problematic characterisations of the scientific community. In particular, SOSR models seem to represent the scientific community as both insulated from society and essentially value-free. While the first issue can be solved—and, indeed, some recent SOSR models have started to incorporate the effects of external influences into their simulated communities—the second problem is more pernicious. From a philosophical point of view, it may appear as if SOSR models support, more or less implicitly, a value-free conception of science. From a technical point of view, it is not even clear how the role of non-epistemic values in scientific reasoning could be incorporated in computer-based SOSR models. That SOSR models are value-free may put into question both their representational adequacy and their normative ambitions.

In the end, it will be argued that examining the limits of SOSR models can in turn shed a new light on the debate about the value-ladenness of science. The social epistemology of science is currently split into two separate and non-communicating branches, one developing formal computer-based SOSR models, the other proposing new descriptive and normative ideals of the moral and political dimension of the scientific community. While the first branch fails to capture the value-ladenness of science, the second fails to address how different individual members of a scientific community may be driven by different non-epistemic values in different ways. This article is a first step to reconcile the two branches of the social epistemology of science.

2 On the current debate on SOSR models and their function

The pioneering work on SOSR modelling was made by Kitcher (1990), who developed an analytical framework borrowed from economics in which self-interested agents seek to maximize their utility. More recent approaches in SOSR modelling make use of Agent-Based Models (ABM), computer programs in which social groups are represented through a number of individual agents, the rules for their behaviour, and their environment. The simulations run on ABMs explain how,

through the interactions of the individual agents, macro-phenomena emerge at the group level (Axelrod, 1997; Gilbert & Terna, 2000). ABMs have been used by social scientists for a long time. Social epistemologists have applied this tool to study the social aspects of science.

At least in the intentions of their developers, SOSR models can be used for advancing suggestions on how to design optimal research teams. Kitcher (1990), for instance, regarded his own model as providing the basis for policy recommendations: “How do we best design social institutions for the advancement of learning? The philosophers have ignored the social structure of science. The point, however, is to change it” (Kitcher, 1990: 22). Similarly, Zollman (2007) uses the results of his simulations to suggest which communicative structures are more preferable than others in some circumstances. Weisberg and Muldoon (2009) also use the results of the simulations run on their model to make suggestions about which populations of agents best distribute cognitive labor and is therefore preferable to other scenarios. SOSR models have had such a normative ambition since their inception, but recently their interest in providing actual policy suggestions has become more explicit (Petrovich & Viola, 2018). Some of them, for example, are developed with the aim of informing resource allocation in scientific research (Avin, 2018; Kummerfeld & Zollman, 2016).

The proliferation of SOSR formal models has led to a growing philosophical literature. A great part of this literature focuses on technical issues, such as the robustness of these models, that is, their capability of producing outputs which remain consistent even when some of their input parameters are modified (see, among the others, Alexander et al., 2015; Rosenstock et al., 2017; Frey & Šešelja, 2019). The analysis of the robustness of particular SOSR models motivates discussions about their ability to inform real-world scientific communities. Apart from focussing on some technical aspects of some of these models in particular, however, recently some philosophers have begun to discuss SOSR models and their representational adequacy in more general terms. What has begun to be questioned, in other words, is not the representational adequacy of one specific SOSR model or another but, rather, the viability of the formal modelling of the scientific community as a philosophical project.

Martini and Fernández Pinto (2017) argue that, contrary to the ABMs employed in the social sciences, SOSR formal models fail to connect with empirical data. As Martini and Fernández Pinto point out, modelling involves (at least) three steps: first, the identification of a problem and the selection of a theoretical framework to tackle it; second, the construction of a model in the light of some specific assumptions; third, the establishment of a connection between the model and the empirical data. In the social sciences, the last step is taken by following a number of strategies, such as parametrization and model validation. In SOSR modelling, however, no step is taken to bridge the gap between models and empirical data. The failure of a model to connect with empirical data is the failure of that model to adequately represent its target system. Therefore, in Martini’s and Fernández Pinto’s view, if SOSR models want to have any function at all, it is crucial for them to take up the ‘empirical challenge’. Similarly, Thicke (2020) argues that if SOSR models are descriptively inadequate, not only because of some of their problematic assumptions (second step of the modelling process), but mainly

because they do not connect in any way with the empirical data (third step), then they ought not be used to support normative claims about the social organisation of science. Bedessem (2019) also points out that, since they do not capture some important aspects of scientific research, such as pluralism and the co-existence of interlocking sub-objectives, SOSR models end up misrepresenting scientific communities and may lead to problematic policy advice.

While some philosophers have questioned the descriptive adequacy and the normative function of SOSR models, others have questioned *whether* their actual function is to provide adequate descriptions and to make normative claims. Frey and Šešelja (2018) suggest that these models should not be viewed as representations in need of empirical validation, but as formal tools which supplement the empirical basis used within some debates in the history and philosophy of science. Šešelja (2020), furthermore, claims that SOSR formal models are ‘abstractions’, the function of which is neither representational nor explanatory but, rather, ‘exploratory’. She argues that the epistemic function of these models is not to provide plausible explanations of the dynamics of actual epistemic communities, but to explore philosophical questions such as those concerning the theories of scientific rationality.

The philosophical and theoretical function of SOSR models is also advocated by Aydinonat et al. (2020), who feel that focussing only on the model-target dyadic relation is limiting. In their view, SOSR models are some of many possible argumentative devices which can be used to defend and reinforce philosophical positions. Following this reasoning, an SOSR model is to be considered successful if it generates further arguments. For example, the Weisberg-Muldoon epistemic landscape *together* with its subsequent versions, which were developed by criticising and then modifying some of its original assumptions (see, for example, Alexander et al., 2015; Thoma, 2015; Pöyhönen, 2017; Hessen, 2019), are all part of the same ‘argumentative landscape’, each new version of the model corresponding to an argumentative move. As such, and despite the limitations of the initial version, the Weisberg-Muldoon epistemic landscape is a successful SOSR model. Considering SOSR models as possessing a philosophical and argumentative function would rescue them from some of the criticisms about their limitations: if their function is *not* to provide an adequate representation of actual scientific communities, then the fact that they are not empirically validated is not a problem anymore. By doing so, however, the idea that SOSR models can be used to justify science-policy recommendations will have to be given up, or at least weakened, despite the fact that such an idea is explicitly held by some of their developers.

Even accepting the idea that SOSR models are argumentative and philosophical tools, there are still a lot of questions to be answered. What is, exactly, the ‘philosophical argument’ these models are making? Which image of science do they support? To answer these questions, I will uncover the philosophical underpinnings shared by many of the current SOSR formal models.

3 The Kuhnian project

The aim of this section is to make explicit what philosophical problems motivated the introduction of SOSR models. To do so, I propose to interpret them as articulations of the *Kuhnian project* in philosophy of science. By ‘Kuhnian project’ I mean something different from the famous model of scientific change developed by Kuhn in *The Structure of Scientific Revolutions* (1962) and in other writings. Regardless of the tenability of that model, it is possible to appreciate Kuhn for establishing a new way of practicing philosophy in which normative claims about science are grounded on a descriptive account of the dynamics of the scientific community. Kuhn did not regard scientific knowledge as the product of an isolated individual genius: in his view, the community of specialists is “the agent of science” (Hoyningen-Huene, 1993: 8, 65, 82). This is why his philosophy should be regarded as one of the first attempts at a *social epistemology of science* (Wray, 2011). As De Langhe (2013) suggests, the recent approaches to the study of the scientific community, such as computer simulations, ABMs, network systems, and big data analysis, can all be regarded as an expansion of Kuhn’s original philosophical method. As I will show, current SOSR models are continuous with the ‘Kuhnian project’ not only because their target is the scientific community, but also because they attempt to solve some of the problems originally introduced by Kuhn.

Kuhn’s interest in the scientific community predates the publication of *Structure*. Based on his own personal experience as a scientist, as well as on a number of philosophical considerations, he provided a general description of the scientific community as pervaded by an *essential tension* (Kuhn, 1959). In Kuhn’s view, the scientific community is constituted by a majority of scientists who have a ‘dogmatic’ attitude towards the dominant theory, methodology, or research project. At the same time, the community also includes a minority of ‘divergent’ and iconoclastic scientists, ready to challenge the dominant research tradition and to explore innovative approaches. Both parties perform vital tasks for science: the dogmatic scientists develop, articulate, and apply the established knowledge, while the fewer divergent thinkers guarantees a ‘reserve tank’ of alternative theories and methods to be used in case of crisis.

Kuhn’s description of the essential tension supports a normative argument about ‘risk-spreading’ in science (D’Agostino, 2010). If every member of the scientific community bets on the same theory or on the same problem-solving strategy, then the group as a whole would risk too much: the theory everybody has endorsed may lose the bet after all, while potentially successful but under-confirmed research strategies would have remained underdeveloped or even completely abandoned. By allowing a minority of divergent scientists to work on alternative theories and methods, the community as a whole hedges its bets and spreads the risk of error. This means that not only the scientific community is pervaded by an essential tension, but it also *ought* to be so.

What is rational for a community as a whole, however, is not necessarily rational for its individual members. For an individual, it is not rational to endorse

a less confirmed theory. In order to preserve the essential tension, however, some scientists ought to do precisely so. As Kitcher observes, there is a “mismatch between the demands of individual rationality and those of collective (or community) rationality” (Kitcher, 1990: 6). He therefore develops his ‘Marginal Contribution-Reward’ (MCR) analytical model to show why it is rational for self-interested scientists *not* to choose the more confirmed theory or the more promising method. In the MCR model, the scientific community has an ‘objective’ (for example, finding a piece of ‘significant truth’) and a pre-established number of available strategies to reach it (for example, theories, hypotheses, or models). In this model, every agent knows both the amount of available strategies and the probability of their success. They all know that, for example, one theory is more empirically confirmed than the others and, therefore, it has a higher probability of success. In short, the agents in this model have all access to the same information and they all believe in the same things. Their choice, however, is not determined by the probabilities of success of the available theories or methods. The rational agents of Kitcher’s model, in fact, are motivated by the probability of making a ‘profit’ (for example, making a valuable contribution that may boost their career). In other words, they use the probability of success of the available strategies to reach the objective in order to calculate the probability of their own personal and professional success. In some cases, the probability of making a profit by working on a less confirmed theory (which is endorsed by fewer scientists and which may be developed in new and surprising ways) is higher than the probability of making a profit by working on the most probable theory (which has already been chosen, and worked on, by the majority of scientists). In short, by choosing a research strategy with a lower probability of success, some scientists may actually increase their ‘expected utility’ in terms of career advancement, visibility, and prestige. The self-interested choices of rational individuals allows the community as a whole to hedge its overall bets. In this way, the cognitive labor necessary to reach the epistemic ends of the community as a whole is divided across different individuals and the Kuhnian essential tension is preserved.

While Kitcher’s MCR framework complements Kuhn’s risk-spreading argument, the more recent computer-based SOSR models modify some of its problematic assumptions. One of them is that Kitcher’s agents possess a complete knowledge of the current state of research, of its internal division in discrete research strategies, and of the probability of their success. Zollman (2007, 2010) overcomes this difficulty: instead of being all-knowing credit-seekers, the agents of his ‘epistemic network’ can and do learn from one another, and have their preferences modified accordingly. In this way, Zollman’s model provides a refined account of how scientists choose their projects.

In both Kitcher’s and Zollman’s models, however, agents do not show important individual differences: they are either rational agents calculating marginal contributions, in one case, or Bayesian agents learning from one another and with experience, in the other. The ‘epistemic landscape’, first developed by Weisberg and Muldoon (2009), overcomes this issue. This model represents a research field as a landscape containing two ‘peaks’. Agents move from one patch of the landscape to the other in an attempt to find the best path to get to the top of the peaks. Different

kinds of agents adopt different exploratory strategies: ‘mavericks’ are adventurous and always choose unknown paths, while ‘followers’ do not take risks and prefer to take the paths others have already opened. In this way, the Weisberg–Muldoon model provides a more sophisticated and refined view of the distribution of cognitive labor.

Finally, all the models discussed so far take for granted the availability of scientific theories but cannot explain where *new* theories come from. De Langhe (2014), who explicitly acknowledges his debts towards Kuhn, develops a ‘unified’ ABM which shows how the distribution of cognitive labor leads to both the exploitation of existing alternatives and to the endogenous generation of new theories within the scientific community.

The SOSR models briefly discussed in this paper lay on an ideal continuum originated by Kuhn, with each new model modifying some of the assumptions of the previous ones, in order to overcome their limitations. To use the terminology of Aydinonat et al. (2020), all these models belong to the same ‘argumentative landscape’ (see *Table 1*).

What kind of argument are these models making? And what image of science they end up supporting? Some philosophers have criticised Kuhn for developing an image of the scientific community as isolated from the rest of society and as ‘value-free’, that is driven by epistemic values only. Apart from the family of philosophical problems they work on, it is worth considering whether current SOSR models also inherit from Kuhn such a contested image of the scientific community. Before discussing whether this is indeed the case, I will briefly explain what the so-called ‘value-free ideal of science’ amounts to.

4 ‘Value-free’ vs. ‘value-laden’ ideals of science

For the so-called ‘value-free ideal’ (VFI), the justification of scientific claims, such as the acceptance or the rejection of a hypothesis, must not be influenced by social, moral or political values. Occasional remarks made by scientists and philosophers provide some general formulations of VFI. Max Weber (1917), for example, distinguished between ‘fact questions’ and ‘value questions’, arguing that science ought to answer only the former, while Henri Poincaré (1920) stated that science and ethics never meet. The main reason for endorsing VFI is a concern for the epistemic authority of science: value-judgments, in fact, could threaten the impartiality and objectivity of science.

Two important specifications are in order. First, VFI does not claim that every kind of value should be eliminated from every stage of scientific activity. Sometimes evidence alone is not sufficient to guide scientists in choosing between two (or more) theories. In such cases, the choice is guided by so-called ‘epistemic values’, such as simplicity, coherence, predictive power, and so on (Kuhn, 1977; Laudan, 1984; McMullin, 1983). VFI does not prescribe to epurate science from epistemic values, but rejects the idea that theory appraisal can be legitimately influenced by *non-epistemic* values, such as social, moral, or political values. The epistemic/non-epistemic values distinction is therefore at the heart of VFI.

Table 1 The development of the 'Kuhnian project'

	Problem	Name of the Model	Type of Model
Kuhn	Showing why it is rational for the scientific community as a whole to spread the risk among individual scientists	Essential Tension	Generalisation from personal experience
Kitcher/Strevens	Resolving the tension between collective and individual rationality	Marginal Contribution Model	Analytical framework
Zollman	Explaining the mechanisms of information transmission within the scientific community	Epistemic Network	Computer-based network analysis
Weisberg & Muldoon	Showing the optimal division of cognitive labor for epistemic progress	Epistemic Landscape	ABM
DeLanghe	Showing the optimal division of cognitive labor for balancing the exploitation of existing theories and the exploration of new theories	Unified model	ABM

Second, VFI recognizes that non-epistemic values can still play a legitimate ‘external’ role in science. Supporters of VFI accept that social and political worries may influence science at the agenda-setting stage, by directing researchers’ attention towards some areas of investigation. They also accept that integrity rules are in place for non-epistemic reasons: perhaps potentially harmful experimentation on human subjects could lead to interesting discoveries, but some research methodologies are simply rejected as unethical. Once the research objectives have been fixed and the research methodology has been approved, VFI states, then the ‘internal’ working of scientific reasoning ought to unfold free from the influence of non-epistemic values.

Several philosophers have questioned the feasibility of VFI. To begin with, the lack of precise criteria for establishing which values should count as ‘epistemic’, and why, makes their differences with the non-epistemic values tenuous, context-dependent, and vague. As a consequence, it is argued that VFI rests on a shaky ground (Anderson, 1995, 2004; Longino, 1990, 2002; Rooney, 1992).

Another argument against VFI is that of *inductive risk*. As discussed by Churchman (1948) and Rudner (1953), the risk of accepting a false hypothesis or of rejecting a true hypothesis cannot be evaluated solely in probabilistic terms. To decide what is ‘worth the risk’ involves a type of value judgment that goes beyond statistical reasoning, especially if error may have potentially harmful consequences. Douglas (2009) extends this argument by showing that scientists face inductive risk at many stages of research, and not only when they have to accept or reject a hypothesis. In choosing the level of statistical significance, in gathering and characterising ambiguous evidence, and in evaluating the support that the interpreted evidence gives to a hypothesis, scientists are faced with the risk of making potentially harmful errors. Since uncertainty is involved in many aspects of scientific research, Douglas concludes that non-epistemic values play a legitimate internal role in scientific reasoning (see also Biddle, 2013; John, 2015).

Both VFI and the arguments against it considered so far are about whether non-epistemic values play a legitimate role in guiding scientists towards epistemic ends. As Elliott and McKaughan (2014) argue, however, apart from its main epistemic goals (such as, the discovery of truths about the world), scientific research has a number of context-dependent ‘sub-goals’. Scientific theories or models are not evaluated only with respect to their fit with the world, but also on the basis of the needs of their users. In some cases, scientists may favour non-epistemic values, such as the ability to generate rapid conclusions, over epistemic values, such as accuracy. Brown (2013, 2020) also develops a number of arguments to demonstrate that sometimes non-epistemic values have a priority over epistemic values.

As it appears from this brief discussion, instead of constituting a single position, the ‘value-laden ideal’ (VLI) is composed of a number of arguments against VFI. Supporters of VLI may have different ideas about whether non-epistemic values can be distinguished from epistemic values, whether their action is ‘direct’ or ‘indirect’ (see Elliott, 2011, 2013), or whether epistemic values should always be prioritized. Also, it is not entirely clear what kind of normative conclusion VLI supports. While it is possible to accept that non-epistemic values may play an important internal role in scientific reasoning, it is not clear whether such values *always* have a positive effect in science. A principled way to distinguish between ‘desirable’ and

‘undesirable’ effects of non-epistemic values in scientific reasoning, however, is yet to be agreed upon.

It must be said that some philosophers defend VFI. Some argue, for instance, that scientists do not have to make decisions and, therefore, they do not face the problem of inductive risks. In this view, scientists’ job is just to communicate the level of uncertainty to policy makers, who are the ones with the responsibility of making decisions (Betz, 2013; Hudson, 2016). The separation between value-free scientists and policy-makers is not always tenable. In some scientific and technological fields, such as in so-called ‘innovation and development’ research, scientists find themselves working in projects which may alter society in unpredictable ways. It is in these fields that scientists, even though they are not policy-makers, often have to assess the potential impact of their research. The risks, in these cases, do not involve only the chance of accepting a false hypothesis or of rejecting a true hypothesis. The risks, rather, have to do with the transformative power of science and technology which may solve some problems, but also create new and unexpected ones (von Schomberg & Block, 2018). On the basis of these considerations, several governance approaches have been developed with the aim of fostering and institutionalising ethical reflection within the scientific community. Examples of such approaches are: the ‘Ethical, Legal and Social Implications’ framework (Balmer et al., 2016; Fisher, 2005); ‘Technology Assessment’ (TA) (Grunwald, 1999, 2018; Rip et al., 1995; Schot & Rip, 1997); ‘Responsible Research and Innovation’ (von Schomberg, 2013). It appears, therefore, that VFI is not just contested on philosophical grounds. That scientists themselves ought to make value judgments has become a requirement of many policies regulating funding allocations.

The aim of this article is not to defend either VFI or VLI. For the purposes of the present work, it is enough to explain that these two different ideals of science exist. It remains to be seen which of them is supported by, or implied by, SOSR models.

5 SOSR models and the ideals of science

5.1 Kuhn’s model

In Kuhn’s view, science makes progress by solving the problems dictated by a dominant paradigm, which sets the research agenda and may isolate the scientific community from the rest of society: “A paradigm [...] can even insulate the [scientific] community from those socially important problems that are not reducible to the puzzle form” (Kuhn 1996: 37).

Towards the end of *Structure*, Kuhn even suggests that scientific progress may depend on the relative isolation of the scientific community from the rest of society:

“Some of these [aspects of scientific progress] are consequences of the unparalleled insulation of mature scientific communities from the demands of the laity and of everyday life. That insulation has never been complete - we are now discussing matters of degree. Nevertheless, there are no other professional communities in which individual creative work is so exclusively addressed to

and evaluated by other members of the profession [...]. Even more important, the insulation of the scientific community from society permits the individual scientist to concentrate his attention upon problems that he has good reason to believe he will be able to solve". (Kuhn 1996: 164).

For Douglas (2009: ch. 3), Kuhn's insular conception of the scientific community contributed to the popularity of VFI among scientists and philosophers after the Second World War. The relation between Kuhn's insular image of the scientific community and VFI, however, might not be as straightforward as Douglas supposes. Kuhn does not claim that scientists ought to avoid every socially important problem, but only those that are not reducible to a tractable form. Not every socially relevant problem may have a scientific solution after all. Kuhn's claims, therefore, are not necessarily a defense of VFI, but rather the recognition of the limits of science. Nor can his claims automatically be regarded as an argument in favour of the unaccountability of science. Kuhn does not say that science ought to be insulated from all possible criticisms, but only that criticisms should be "competent enough to be taken seriously by scientists and [...] constructive enough to offer a viable alternative" (Mladenović, 2017: 134). In and by itself, describing the scientific community as isolated from society does not necessarily imply VFI. In principle, it is possible to conceive a scientific community which is *both* separated from the most immediate but not always tractable demands of society *and* internally driven by moral and social values.

The problem is that Kuhn also seems to characterise the scientific community as being internally driven by epistemic values only. For Kuhn (1977), what guides scientists in theory choice and theory appraisal is a set of epistemic values, such as accuracy, simplicity, consistency, scope, and fertility. In his view, the list of epistemic values does not change through time: they are the essential criteria for scientificity, in the sense that a theory, to be considered scientific, must display them. What changes is the way in which these values are ranked and applied. Some scientists may consider accuracy more important than fertility: therefore, they would prefer a different theory than the one chosen by those who value fertility over accuracy. The application and ranking of epistemic values, in other words, is not algorithmic: this has led some philosophers to talk about 'Kuhnian-underdetermination' (Carrier, 2008).

Non-epistemic values could offer a possible way out to this problem. Moral and social considerations may tell scientists in which cases, for example, accuracy should be preferred to fertility or vice versa. Kuhn describes the scientific enterprise as being driven by epistemic values only, but the very way in which he characterizes them may leave the door open to non-epistemic values. This, however, is a road that he himself did not take: in none of his works Kuhn discusses how non-epistemic values could solve the problem posed by the non-algorithmic nature of epistemic values. As a matter of fact, the discussions about the role of non-epistemic values in prioritizing and applying epistemic values are usually developed as a corrective to some aspects of Kuhn's philosophy. For example, even though she inherits from him an interest for the social structure of science, and although she does not want to replace Kuhn's list of values, Longino (1990, 2002) disagrees with the very idea

of context-independent and invariable criteria of scientificity. Others, as mentioned in the previous section, question the very distinction between epistemic and non-epistemic values (Rooney, 1992).

What emerges from many of Kuhn's published works is an image of the scientific community as insulated from the rest of society and as driven only by purely epistemic values. Kuhn never formulated any arguments against VFI, nor did he explicitly discuss the role of non-epistemic values in scientific reasoning. Therefore, his position can be regarded as being at least compatible with VFI. Since it is based on generalizations and it did not rely on a strict formalism, it is not too difficult to amend or expand upon Kuhn's descriptive/normative model of the scientific community, as some philosophers have done. It remains to be seen whether subsequent formal SOSR models accepted the value-free image of the scientific community implied in many of Kuhn's works, instead of problematizing or correcting it.

5.2 Kitcher's model

Kitcher's model relies on an analytical framework borrowed from classical economics. The behaviour of its profit-seeking agents follows the principles of Rational Choice Theory and, in the long run, an 'invisible hand' will bring the whole system to a state of equilibrium. This model has been criticised for simply presupposing the validity of the economics framework, which instead would deserve further philosophical reflection (Hands, 1995, 1997). Moreover, 'invisible hands' mechanisms may just fail to explain the success of science (Wray, 2000). Apart from these considerations, and for the purposes of this article, it should be wondered which picture of the scientific community the MCR model provides.

Kitcher (1990) makes clear that the 'objective' represented in his model is a piece of 'significant truth': something that is not only true, but that it is also considered of some importance in a given context. His example is that of the discovery of a molecule to cure a disease. Clearly, the discovery of such a molecule is socially relevant. This, however, does not make the model 'value-laden'. As explained in the previous section, supporters of VFI do not reject the idea that moral or social considerations may influence, or even determine, the direction of scientific research. Simply defining the objective of the scientific community represented in the model as 'significant', therefore, is still compatible with VFI.

When it comes to the internal working of the scientific community, in Kitcher's idealized model scientists are driven by a mix of epistemic aims (i.e., finding the right solution) and selfish interests (i.e., increasing utility). His model does not account for the role of non-epistemic values in scientists' decisions. It is possible to argue that the MCR model not only is compatible with VFI, but it actually supports an image of the scientific community as being indeed value-free.

As already discussed, many philosophers argue that non-epistemic values play an internal role in many steps of scientific research. Individual differences in scientists' application and ranking of non-epistemic values may impact the distribution of cognitive labor Kitcher talks about. Following an example provided by Elliott (2017: ch. 4), let's assume that the aim of a project in the field of agricultural science is to find

a way to produce more food in poor countries suffering from hunger issues. Let's assume that there are two possible ways to reach such an objective: the first consists in investigating biotechnological venues for the production of genetically modified food, the second in analysing the characteristics of the local land in order to develop and implement better agricultural methods. Finally, let's assume, as Kitcher would, that all the scientists involved in the pursuit of that objective know in advance that the probability of success of the first approach (biotechnological innovation) is higher than the probability of success of the second approach (study of local agricultural land). If scientists behaved as in Kitcher's model, then the majority of them would choose the first method, while for a minority of them it would be more convenient to choose the second. The credit-seeking individual behavior would guarantee an optimal distribution of the cognitive labor and maintain the essential tension: while the majority of scientists will choose the most probable approach, a minority will find it convenient to choose the less probable approach, which may nevertheless become useful in the future. Things appear different if one takes into consideration that scientists are not driven only by epistemic values and selfish motivations, but also by non-epistemic values.

Let's assume that, in our example, the first approach, which is also the most probable in terms of success in reaching the objective, is also the most profitable in financial terms. Innovative research in biotechnology, in fact, may lead to the development of new techniques which could be patented and sold to biotech companies. On the one hand, scientists who do *not* choose the most probable method may not be motivated by the probability of future rewards (that is, they may not be driven by the possibility of making a relevant contribution in the less crowded and less competitive arena). Rather, these scientists' choice could be motivated by the desire of avoiding methods which they perceive as leading to environmental issues, such as the impoverishment of the local terroir. For them, in short, the first method is simply too risky, if not plain dangerous for the local communities. On the other hand, those who choose the first method may not be motivated by its higher probability of success, or by the possibility of a financial return, or by a mixture of the two. If not a definitive solution, genetically modified food can still offer a very quick fix to the problem of hunger. For the scientists choosing the first method, therefore, what is really too risky is waiting too long in a context in which people's actual life is at peril.

The interplay between epistemic and non-epistemic considerations may lead to a different distribution of cognitive labor than the one represented in Kitcher's model. The majority of scientists may end up choosing the less probable method of research if the method with the highest probability of success is associated with a high risk of potentially harmful consequences. The method with a lower probability of success, however, may also involve potential harm. Depending on what is considered more risky and potentially dangerous for society, different 'distributions' may emerge. Deciding which one is the optimal one becomes a less straightforward task.

Incidentally, in recent years Kitcher has become one of the most prominent philosophers defending the necessity of a value-laden science, driven by moral and democratic values (Kitcher, 2001, 2011). However, he has never seen how such values may alter the distribution of cognitive labor of his own model. Kitcher was

actually aware of the limitations of the MCR formal model, and auspicated that they would have been overcome through the inclusion of non-epistemic factors (other than selfish interests):

“appealing to human ambition is only the beginning of the story. Other psychological mechanisms might bring scientists closer to the [optimal distribution of cognitive labor] than they would otherwise have been. Not only many vices from greed to fraud play a constructive role, but community ends may be furthered by more salubrious traits. Perseverance, personal investment, personal and national loyalties, and devotion to political causes may, on occasion, help to close a [community-individual] discrepancy” (Kitcher, 1990: 18).

As it stands now, however, Kitcher’s model does not account for the role of social, moral and political values in scientific reasoning and, therefore, it ends up supporting a rather value-free image of the scientific community.

5.3 Computer-based SOSR models

Many computer-based SOSR models provide a representation of the scientific community as a complex, self-regulated, and ultimately ‘closed’ system. Such a closedness is a rather strong idealisation: actual social groups, including the scientific community, are influenced by a larger social context; often, social groups and their wider context co-evolve together. Current SOSR models also rely on other dubious assumptions, for instance that the number of agents within a scientific community is constant.

On the one hand, it could be argued that many of these simplifying assumptions are justified: SOSR models are, after all, idealisations with a limited scope. On the other hand, there have been some attempts to relax or to modify at least some of these simplifying assumptions. Holman and Bruner (2017), for example, have developed an epistemic network which takes into consideration the influence of industrial funding mechanisms in agent’s decisions. In the same way in which an insular representation of the scientific community does not necessarily imply a value-free image of science, however, representing the influence of some external factors on the scientific community does not necessarily make a model able to capture the value-ladenness of science. As already discussed, VFI does not deny the external influence of non-epistemic factors on science. What VFI denies is that non-epistemic factors play a legitimate internal role in scientific reasoning. Scientists influenced by, or even corrupted by, industrial funding mechanisms are not legitimately using non-epistemic values in scientific reasoning; rather, they are not reasoning correctly. Holman’s and Bruner’s network, in other words, represents how external factors may pollute scientific reasoning, not how proper scientific reasoning is value-laden.

In another model, Holman and Bruner (2015) represent how the false or unreliable information produced by biased agents spreads throughout a network. Clearly, accepting false theories or bad evidence may have very important social and ethical consequences. Modelling how error or false information spread through the network, once again, is far from capturing the legitimate internal role of non-epistemic values

in scientific research. In this model, in fact, agents are either ‘genuine truth-seekers’ or ‘intransigently biased’. Nothing in this model can account for how non-epistemic values guide scientists in assessing theories in the face of uncertainty.

Epistemic landscape models also seem to provide a rather value-free image of the scientific community. Similarly to Kitcher’s framework, these models presuppose the existence of pre-defined objectives (‘peaks’) and of various strategies for reaching it (‘paths’). Weisberg and Muldoon (2009), like Kitcher, define the objective as a piece of ‘significant truth’ that, as such, is relevant and useful to particular groups of people with practical interests. Some would say that representing agents as seekers of significant truths means to represent how scientists work for the ‘common good’. As in the case of Kitcher, however, the objective of the model is predetermined, its ‘significance’ and social relevance are exogenously given: this is all consistent with VFI. *After* such a significant objective has been fixed, in fact, there is no place for non-epistemic values in the choices made by the agents.

As argued by Bedessem (2019), this kind of SOSR models presuppose that the members of the scientific community pursue only the pre-determined significant objectives, represented as discrete parts of the landscape. Actual scientific research, however, pursues several interlocking sub-objectives. To discover an important molecule to cure a disease (main objective), for example, scientists may have to develop new techniques, design new experiments, run new tests (sub-objectives). The significance of such sub-objectives is not just a function of their getting closer to the main objective: some of them may become significant for the pursuit of *other* objectives. A newly developed technique, in fact, may fail to reach the original objective but may nevertheless open up a new line of significant research. Following on Bedessem’s criticism, and as discussed by Elliott and McKaughan (2014), the significance of the sub-goals of science is context-dependent and not determined on purely epistemic grounds. Sometimes scientists may deliberately choose less accurate models if they produce faster results and if they lessen social costs. In other words, scientists may employ non-epistemic considerations for choosing between or prioritizing different sub-objectives.

While ‘epistemic significance’ is an intrinsic feature of the landscape (higher patches are more significant than lower patches), the model does not represent how agents choose paths on the basis of non-epistemic factors. This means that, in this SOSR model, agents’ decisions are made on purely epistemic grounds. Even though ‘mavericks’ and ‘followers’ have different exploratory attitudes, in fact, they follow the same fundamental rule: to move from lower to higher patches. Moving upward is the fundamental criterion agents use to explore the landscape. If they end up in a lower patch, they simply go around trying to move up again. Nothing in the model indicates how agents choose between two different but equally ascending paths, nor does the model represent how some agents may sometimes prefer to move towards slightly lower patches in cases in which non-epistemic considerations take priority over epistemic values.

It is worth noticing that SOSR models are completely silent about the ‘moral profile’ of their agents. In epistemic landscape models, for instance, there is no ‘rule’ that agents follow to deal with the potential impact of choosing the wrong path. Mavericks are very quick in exploring unknown parts of the landscape, but there is

no real consequence for their possible errors or risky decisions. In a real scientific community, mavericks' behaviour would be considered far too reckless. One way to correct this aspect of the model would be to provide its agents with the ability to ponder the risk of error in order to avoid potentially harmful consequences. Such a reflexivity may slow down 'responsible mavericks'. As a consequence, a different mavericks/followers ratio than the one in Weisberg's and Muldoon's simulations will be optimal.

SOSR models do not account for the role of non-epistemic values in scientific reasoning. Moreover, it is not entirely clear how computer-based SOSR models could include this aspect of science in their representations of the scientific community.

6 Can SOSR models represent the value-ladenness of science?

Computer-based SOSR models represent scientists as rule-following agents moving towards epistemic objectives. To capture the value-ladenness of science, these models may provide agents of the models with some new rules to show how they make value-judgements and critical decisions. Such a task, however, poses some problems.

As already mentioned, some of the most recent SOSR models are computer based Agent-Based Models (ABMs) like those which are widely employed in the social sciences. ABMs are useful tools for the analysis of problematic social phenomena which, for obvious reasons, cannot be explored empirically. One of the first and most known ABM, for example, provides a formal representation of the group dynamics leading to the phenomenon of ethnic segregation (Schelling, 1971). More recent models, just to mention a few, represent ideal scenarios in which criminality can be successfully inhibited (Birks et al., 2012), or in which the group dynamics may lead to the escalation of radicalization (Neumann, 2014). In short, ABMs are often employed by social scientists to represent and frame several social phenomena with clear moral overtones: these models are, in a sense, 'morally charged'.

It is important to stress that what the ABMs used in the social sciences do is to model individual preferences and actions leading to large-scale group phenomena which the individual agents did not necessarily intend. To make an example: Schelling's famous model of the production of social segregation represents how, as a result of individual preferences about the choice of the neighborhood, a social group may end up segregating sub-groups belonging to ethnic minorities, even though none of the agents is programmed to have 'racist preferences'. In Schelling's model, residential segregation of ethnic minorities emerges from the complex and re-iterated interactions of virtually non-racist agents: the macro-phenomenon is an unintended consequence of individual actions. Although used as a tool to understand morally charged social macro-phenomena, Schelling's model is not in itself a model of the moral makeup of its individual agents. The claim that science is not value-free, by contrast, amounts to the claim that scientists apply non-epistemic values to make critical decisions and to act responsibly. The social responsibility of science is not the spontaneous or even accidental property emerging from the interaction of a-moral individuals. Rather, science is socially responsible inasmuch as scientists act ethically. This means that value-ladenness can be

properly integrated in SOSR formal models if such models could represent scientists as moral agents.

The possibility of designing moral agents in computational ABMs is the object of an on-going debate within the AI community. As suggested by Ruvinsky (2008), ‘rights’, ‘liberties’, ‘duties’, and other elements constituting a moral framework can be regarded as parameters which can be implemented in what she calls ‘computational ethics’. Such parameters can be used to model agents holding different ethical standpoints. In this way, it is even possible to represent an idealised society in which agents hold different ethical standpoints. Ruvinsky’s view is purely theoretical, since she does not clarify how the parameters making up ‘moral frameworks’ can be implemented in practice into a computational ABM.

An ABM computer simulation which actually implements an ethical theory for the design of its agents has been developed by Mascaro et al (2010). Their ABM represents an evolving world, in which the interacting agents can reproduce and pass some of their traits and behaviors to the next generation. With this model, they study how phenomena such as altruism or suicide emerge and spread across the idealised evolving world. The programmers have modelled the moral agents by relying on utilitarian theories. In fact, they explicitly claim that utilitarianism is the *only* ethical theory which can be successfully implemented in a computer simulation: “In order for computer simulation studies to be informative about ethics, we must adopt a point of view which allows us to measure the outcomes. Utilities are the natural currency for measuring ethical outcomes. Utilities also support a very natural ethical system, namely utilitarianism, the thesis that what action is best collectively is what action is best. Utilitarianism is, in fact, the only ethical system which allows us to measure the outcomes of computer simulations and judge them as better or worse” (Mascaro et al., 2010: 5). By contrast, other ethical theories, such as deontological ethics or virtue ethics, “depend upon the exact semantics of the deontic principles or the virtues, respectively, and incorporating semantic understanding into artificial life simulation in any kind of sophisticated way requires a prior solution to the problem of natural language understanding” (Mascaro et al., 2010, p. 32).

In order to represent scientists as moral agents in a computer-based SOSR model, it is first necessary to specify which ethical theory can be successfully implemented in an ABM, but also which ethical theory best describes socially responsible scientists. If, on the one hand, utilitarianism seems a good candidate for modelling moral agents in an ABM, on the other hand it is questionable whether responsible scientists are best described by such an ethical theory. Constructing ABM models with moral agents is not impossible in principle. It is still not clear, however, whether an SOSR formal model would be able to *adequately* represent moral scientists. It is also unclear whether such a model would be able to represent the interplay between the ethical and the epistemic dimension: to describe, that is, how and to what extent scientists are driven by *both* epistemic *and* non-epistemic values.

7 Value-free models in social epistemology, social-free debates about the value-ladenness of science?

In this article, SOSR models have been interpreted as the development of the ‘Kuhnian project’ in philosophy of science. Kuhn’s view on the essential tension is at the origin of the subsequent SOSR models, with each new model attempting to solve some of the problematic assumptions of the previous ones. Their increasing formalization, however, has led SOSR models to exacerbate some limitations of Kuhn’s account: namely, his view of the scientific community as both insular and, above all, value-free. SOSR formal models do not represent the internal role of non-epistemic values in scientific reasoning and, in the case of computer-based models, it is also difficult to see how they could provide such a representation. But *should* they?

Some would simply admit that there is more work to do. Constructing artificial models of the scientific community involves a lot of preliminary assumptions, methodological choices, and so on. So far, it has been easier to construct value-free models of science. This does not exclude that, in the future, some new SOSR model will be able to account for the role of non-epistemic values in agents’ choices. The argument of this article would therefore amount to something similar to the ‘empirical challenge’ launched by Martini and Fernández Pinto (2017): a sort of ‘moral challenge’ or, in less grandiose terms, an invitation to include the value-laden dimension of scientific reasoning in future modelling of the agents’ behaviour.

Others, however, would not regard the value-free character of SOSR models as a problem. After all, they could say, formal models in the social epistemology of science, like those employed in the social sciences, are not expected to provide a *complete* picture of their target. It could be argued, therefore, that trying to explicitly represent the role of non-epistemic values in the agents’ underlying methodological choices would do anything given the purpose of these models. As already mentioned in the second section, however, the problem is that it is not always clear what the purpose of these models is supposed to be.

If SOSR models are regarded as providing the basis for advancing policy suggestions about real scientific communities, then the fact that they fail to represent the value-ladenness of their target is clearly a problem. In fact, if the models are descriptively inadequate then they can hardly be used to make prescriptions. Moreover, and as mentioned at the end of the fourth section, current science policy frameworks demand scientists to be socially responsible and to explicitly engage in anticipative reflection about the possible societal implications of their work. It is hard to see how value-free abstract models of the scientific community can be of any use in a policy context which rewards science if it is ‘open’, ‘democratized’ and, ultimately, ‘value-laden’.

If, on the other hand, these models are regarded as philosophical and argumentative tools, as some philosophers claim, then it must be wondered which kind of argument they are making. These models represent an often isolated and value-free scientific community, whose agents are driven only by epistemic reasons. It is therefore difficult not to see them as argumentative tools which can be used to make an argument consistent with, or even in support of, VFI.

It is worth stressing that Kuhn's view on the scientific community not only is at the origin of the development of current SOSR models, but it has also been discussed by those philosophers who, as a reaction, have proposed different models of the scientific community. Normative models of the moral dimension of the scientific community and of its relations with the rest of society, such as those proposed by Longino (1990, 2002), Kitcher (2001, 2011), Kourany (2010), and Solomon (2001), are all elaborated within the so-called social epistemology of science. This means that both the formal and 'value-free' SOSR models and the value-laden images of science are part of the same philosophical tradition. Contemporary social epistemology of science indeed appears fractured into two debates, which develop independently from one another, and which are carried on by two non-communicating groups of philosophers. Emblematic is the case of Philip Kitcher, who has contributed to both debates but without providing a 'unified model' of the distribution of the cognitive labor and of the non-epistemic values within the scientific community (which is ironic, considering how much Kitcher values the virtue of unification).

Reflecting upon some of the limits of the SOSR models contributes to discussing the current state of the social epistemology of science and, also, to highlighting some important limits in the current debates on the value-ladenness of science. If, on the one side of the social epistemology of science, there are SOSR models missing the value-laden dimension of science, on the other side, there are debates about the value-ladenness of science which miss the social dimension of actual scientific research. What is usually argued, within such debates, is that, to say it like Rudner (1953), "the scientist *qua* scientist makes value judgments". The problem is: *which* scientist makes value judgments? The value judgment maker invoked in this kind of debate is *the* scientist: an ideally moral and responsible individual driven by both epistemic and non-epistemic values and who takes the right decisions in order not to hurt society. This ideal moral scientist is as much an inadequate representation of actual scientists as the ideal rational agents of the SOSR models. But while SOSR models were developed to illustrate, and possibly resolve, the tension between individual and collective aims, the debate on the role of non-epistemic values in science has not started yet to frame the problem of the tension between individual moral character and collective responsibility.

The 'moral challenge' to SOSR models and the 'social challenge' to the debate on non-epistemic values in science are, in a sense, complementary. Attempting to bridge these two separate branches of the philosophy of science might lead to more complete representations of the scientific community and, consequently, to an increased understanding of its dynamics.

8 Conclusions

Like Kuhn's original framework, SOSR formal models represent the scientific community as a value-free system. While it is possible to justify such a representation by specifying that it is an idealization or an abstraction, it remains to be seen whether such an idealization or abstraction is either useful or even desirable. It could be argued that, since non-epistemic values do play a crucial internal role in actual

scientific research, the failure to capture the value-ladenness of science makes SOSR models descriptively inadequate, if not philosophically problematic. However, integrating non-epistemic values into the makeup of the agents of the current computer-based SOSR models may be a difficult task. Nevertheless, reflecting upon the limitations of SOSR models to capture the value-ladenness of science also illuminates some limitations in the current discussions about the role of non-epistemic values in science.

Declarations

Ethical approval This is a theoretical/philosophical work, which does not require any ethical approval

Informed consent This is a theoretical/philosophical, no research was made on human subjects

Conflict of interest The author declares no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alexander, J. M., Himmelreich, J., & Thompson, C. (2015). Epistemic landscapes, optimal search, and the division of cognitive labor. *Philosophy of Science*, *82*, 424–453.
- Anderson, E. (1995). Knowledge, human interests, and objectivity in feminist epistemology. *Philosophical Topics*, *23*, 7–58.
- Anderson, E. (2004). Uses of value judgments in science: A general argument, with lessons from a case study of feminist research on divorce. *Hypatia*, *19*, 1–24.
- Avin, S. (2018). Policy considerations for random allocations of research funds. *Roar Transactions: a Journal on Research Policy & Evaluation* 1, <https://riviste.unimi.it/index.php/roars/article/view/8626/9290>. Accessed 1 July 2021.
- Axelrod, R. (1997). *The complexity of cooperation: Agent-based models of competition and collaboration*. Princeton University Press.
- Aydinonat, N. E., Reijula, S., & Ylikoski, P. (2020). Argumentative landscapes: The functions of models in social epistemology. *Synthese*. <https://doi.org/10.1007/s11229-020-02661-9>
- Balmer, A., Calvert, J., Marris, C., Molyneux-Hodgson, S., Frow, E., Kearnes, M., Bulpin, K., Schyfter, P., Mackenzie, A., & Martin, P. (2016). Five rules of thumb for post-ELSI interdisciplinary collaborations. *Journal of Responsible Innovation*, *3*, 73–80.
- Bedessem, B. (2019). The division of cognitive labor: Two missing dimensions of the debate. *European Journal for Philosophy of Science*. <https://doi.org/10.1007/s13194-018-0230-8>
- Betz, G. (2013). In defence of the value free ideal. *European Journal for Philosophy of Science*, *3*, 207–220.
- Biddle, J. (2013). State of the field: Transient underdetermination and values in science. *Studies in History and Philosophy of Science*, *44*, 124–133.

- Birks, D., Townsley, M., & Stewart, A. (2012). Generative explanations of crime: Using simulation to test criminological theory. *Criminology*, *50*, 221–254.
- Brown, M. (2013). Values in science beyond underdetermination and inductive risk. *Philosophy of Science*, *80*, 829–839.
- Brown, M. (2020). *Science and moral imagination: A new ideal for values in science*. University of Pittsburgh Press.
- Carrier, M. (2008). The aim and structure of methodological theory. In L. Soler, H. Sankey, & P. Hoyningen-Huene (Eds.), *Rethinking scientific change and theory comparison: Stabilities, ruptures, incommensurabilities?* (pp. 273–290). Springer.
- Churchman, C. W. (1948). Statistics, pragmatics, induction. *Philosophy of Science*, *15*, 249–268.
- D'Agostino, F. (2010). *Naturalizing epistemology. Thomas Kuhn and the 'Essential Tension.'* Palgrave Macmillan.
- De Langhe, R. (2013). The Kuhnian paradigm. *Topoi*, *32*, 65–73.
- De Langhe, R. (2014). A unified model of the division of cognitive labor. *Philosophy of Science*, *81*, 444–459.
- Douglas, H. (2009). *Science, policy, and the value-free ideal*. University of Pittsburgh Press.
- Elliott, K. (2011). Direct and indirect roles for values in science. *Philosophy of Science* *78*(2), 303–324. <https://doi.org/10.1086/659222>
- Elliott, K. (2013). Douglas on values: From indirect roles to multiple goals. *Studies in History and Philosophy of Science Part A*, *44*(3), 375–383. <https://doi.org/10.1016/j.shpsa.2013.06.003>
- Elliott, K., & McKaughan, J. (2014). Non-epistemic values and the multiple goals of science. *Philosophy of Science*, *81*, 1–21
- Elliott, K. (2017). *A tapestry of values: An introduction to values in science*. Oxford University Press.
- Fisher, E. (2005). Lesson Learned From Ethical, Legal and Social Implications Program (ELSI): Planning Societal Implications Research for the National Nanotechnology Program. *Technology in Society*, *27*, 321–328.
- Frey, D., & Šešelja, D. (2018). What is the epistemic function of highly idealised agent-based models of scientific inquiry? *Philosophy of the Social Sciences*. <https://doi.org/10.1177/0048393118767085>
- Frey, D., & Šešelja, D. (2019). Robustness and idealisations in agent-based models of scientific interactions. *The British Journal for the Philosophy of Science*. <https://doi.org/10.1093/bjps/axy039>
- Gilbert, N., & Terna, P. (2000). How to build and use agent-based models: Alternative and prospects. *Mind & Society*, *1*, 57–72.
- Grunwald, A. (1999). Technology assessment or ethics of technology? *Ethical Perspectives*, *6*, 170–182.
- Grunwald, A. (2018). *Technology assessment in practice and theory*. Routledge.
- Hands, D. (1995). Social epistemology meets the invisible hand: Kitcher on the advancement of science. *Dialogue*, *34*, 605–621.
- Hands, D. (1997). Caveat Emptor: economics and contemporary philosophy of science. *Philosophy of Science*, *64*, ss107–ss116.
- Hessen, R. (2019). The Credit Incentive to be a Maverick. *Studies in History and Philosophy of Science*, *76*, 5–12.
- Holman, B., & Bruner, J. (2015). The problem of intransigently biased agents. *Philosophy of Science*, *82*(5), 956–968. <https://doi.org/10.1086/683344>
- Holman, B., & Bruner, J. (2017). Experimentation by industrial selection. *Philosophy of Science*, *84*, 1008–1019.
- Hoyningen-Huene, P. (1993). *Reconstructing scientific revolutions: Thomas Kuhn's philosophy of science*. The University of Chicago Press.
- Hudson, R. (2016). Why we should not reject the value-free ideal in science. *Perspectives on Science*, *24*, 167–191.
- John, S. (2015). Inductive risk and the contexts of communication. *Synthese*, *192*, 79–96.
- Kitcher, P. (1990). The division of cognitive labor. *Journal of Philosophy*, *87*, 5–22.
- Kitcher, P. (2001). *Science, truth, and democracy*. Oxford University Press.
- Kitcher, P. (2011). *Science in a democratic society*. Prometheus Books.
- Kourany, J. (2010). *Philosophy of science after feminism*. Oxford University Press.
- Kuhn, T. S. (1959). The essential tension: Tradition and innovation in scientific research. Reprinted in Kuhn, T. (Ed.) *The essential tension: selected studies in scientific tradition and change* (1977) (pp. 225–239). University of Chicago Press.
- Kuhn, T. S. (1962 [1996]). *The structure of scientific revolutions*. University of Chicago Press, 3rd edition.

- Kuhn, T. S. (1977). Objectivity, value judgement and theory choice. Reprinted in Kuhn, T. (Ed.). *The essential tension: Selected studies in scientific tradition and change* (pp. 320–339). University of Chicago Press.
- Kummerfeld, E., & Zollman, K. (2016). Conservatism and the scientific state of nature. *The British Journal for the Philosophy of Science*, 67, 1057–1076.
- Laudan, L. (1984). *Science and values*. University of California Press.
- Longino, H. (1990). *Science as social knowledge: Values and objectivity in scientific inquiry*. Princeton University Press.
- Longino, H. (2002). *The fate of knowledge*. Princeton University Press.
- Martini, C., & Fernández Pinto, M. (2017). Modeling the social organization of science: Chasing complexity through simulations. *European Journal for Philosophy of Science*, 7, 221–238.
- Mascaro, S., Korb, K., Nicholson, A., & Woodberry, O. (2010). *Evolving ethics: The new science of good and evil*. Imprint Academic.
- McMullin, E. (1983). Values in Science, in Asquith, P.D. and Nickles, T. (Eds.). *PSA 1982* (Vol. 2, pp. 3–28). Philosophy of Science Association.
- Mladenović, B. (2017). *Kuhn's Legacy. Epistemology, metaphilosophy, and pragmatism*. Columbia University Press.
- Neumann, M. (2014). The escalation of ethnonationalist radicalization. *Social Science Computer Review*, 32, 312–333.
- Petrovich, E., & Viola, M. (2018). Social epistemology at work: from philosophical theory to policy advice. *Roar Transactions: a Journal on Research Policy & Evaluation* 1. <https://riviste.unimi.it/index.php/roars/article/view/9828>. Accessed 1 July 2021.
- Poincaré, H. 1920[1958]. *The Value of Science*. English translation by Halsted, G., Dover Publications.
- Pöyhönen, S. (2017). Value of cognitive diversity in science. *Synthese*, 194, 4519–4540.
- Reijula, S., & Kuorikoski, J. (2019). Modeling epistemic communities. In M. Fricker, P. Graham, D. Henderson, & N. Pederson (Eds.), *The Routledge handbook of social epistemology* (pp. 250–249). Routledge.
- Rip, A., Misa, T., & Schot, J. (Eds.). (1995). *Managing technology in society: The approach of constructive technology assessment*. Pinter.
- Rooney, P. (1992). On values in science: is the epistemic/non-epistemic distinction useful?, *PSA: Proceedings of the biennial meeting of the philosophy of science association* (pp. 13–22).
- Rosenstock, S., Bruner, J., & O'Connor, C. (2017). In epistemic networks, is less really more? *Philosophy of Science*, 84, 234–252.
- Rudner, R. (1953). The scientist *qua* scientist makes value judgments. *Philosophy of Science*, 20, 1–6.
- Ruvinsky, A. (2008). Computational ethics. In M. Quigley (Ed.), *Encyclopedia of information ethics and security* (pp. 76–82). Information Science Reference.
- Schelling, T. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology*, 1, 143–186.
- Schot, J., & Rip, A. (1997). The past and future of constructive technology assessment. *Technological Forecasting and Social Change*, 54, 251–268.
- Solomon, M. (2001). *Social empiricism*. MIT Press.
- Šešelja, D. (2020). Exploring scientific inquiry via agent-based modeling. *Perspectives on Science* (forthcoming): <http://philsci-archive.pitt.edu/17266/>. Retrieved: 20th January 2021.
- Thicke, M. (2020). Evaluating formal models of science. *Journal for General Philosophy of Science*, 51, 315–335.
- Thoma, J. (2015). The epistemic division of labor revisited. *Philosophy of Science*, 82, 454–472.
- von Schomberg, R. (2013). A vision of responsible innovation. In R. Owen, M. Heintz, & J. Bessant (Eds.), *Responsible innovation* (pp. 51–74). Wiley.
- von Schomberg, L., & Blok, V. (2018). The turbulent age of innovation. *Synthese*. <https://doi.org/10.1007/s11229-018-01950-8>
- Weber, M. (1917). The meaning of 'Ethical Neutrality' in sociology and economics. In Shils, E., & Finch, H. (Eds.), *The Methodology of the Social Sciences* (pp. 1–48). Free Press, 1949
- Weisberg, M. (2010). New approaches to the division of cognitive labor. In P. D. Magnus & J. Busch (Eds.), *New waves in philosophy of science* (pp. 250–269). Palgrave Macmillan.
- Weisberg, M., & Muldoon, R. (2009). Epistemic landscapes and the division of cognitive labor. *Philosophy of Science*, 76, 225–252.
- Wray, B. (2000). Invisible hands and the success of science. *Philosophy of Science*, 67, 163–175.
- Wray, B. (2011). *Kuhn's evolutionary social epistemology*. Cambridge University Press.

Zollman, K. (2007). The communication structure of epistemic communities. *Philosophy of Science*, 74, 574–587.

Zollman, K. (2010). The epistemic benefit of transient diversity. *Erkenntnis*, 72, 17–35.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.