

## EVALUATION OF RESEARCH(ERS) AND ITS THREAT TO EPISTEMIC PLURALISMS\*

MARCO VIOLA

Moscow State Pedagogical University, Russian Institute for Advances Studies

### ABSTRACT

*While some form of evaluation has always been employed in science (e.g. peer review, hiring), formal systems of evaluation of research and researchers have recently come to play a more prominent role in many countries because of the adoption of new models of governance. According to such models, the quality of the output of both researchers and their institutions is measured, and issues such as eligibility for tenure or the allocation of public funding to research institutions crucially depends on the outcomes of such measures. However, concerns have been raised over the risk that such evaluation may be threatening epistemic pluralism by penalizing the existent heterodox schools of thought and discouraging the pursuit of new ones. It has been proposed that this may happen because of epistemic bias favouring mainstream research programmes. In this paper, I claim that (1) epistemic pluralism is desirable and should be preserved; (2) formal evaluation exercises may threaten epistemic pluralism because they may be affected by some form of epistemic bias; therefore, (3) to preserve epistemic pluralism, we need some strategy to actively dampen epistemic bias.*

**Keywords:** *Economic Epistemology, Epistemic Pluralism, Research Policy, Research Evaluation*

### 1. A new governance for research

At the end of the last Century, many national research and higher education systems underwent major reforms toward a new style of governance, named *steering at a distance*. According to his inventor, it involves

a departure from classical government steering by means of legislation, prohibitions and regulations and a move towards the autonomy and the delegation of responsibilities to institutions of higher education [...] Under the new system, intervention is restricted to *ex-post facto* adjustments based on quality assessments of results (Kickert 1995, 135).

This style is inspired by the ideals of *New Public Management*, that prescribes to reproduce quasi-market condition to prompt market-like competition in public institutions. As a consequence of this shift, and in the name of accountability and efficiency of public expenditure, many countries implemented large scale and centralized *ex-post* research evaluation systems (Whitley and Gläser 2007). Rather than merely portraying the state of research in a given country, these evaluation systems significantly affect it, especially (but not exclusively) in the many countries where public funding is based upon their results (Hicks 2012). Furthermore, some countries implemented centralized and formalized procedures to regulate hiring. In the remainder of this article, by speaking about *formal evaluation* I will indicate either these kinds of large scale exercises.

## 2. The unwelcoming reception of formal evaluation in Italy

In Italy, formal evaluation has been introduced later than in most Western countries. Notably, a National Agency for the Evaluation of Research and University System (ANVUR) was founded in 2010, which designed both a centralized system to regulate academic recruitment, i.e. the *Abilitazione Scientifica Nazionale* (ASN),<sup>1</sup> and a nation-wide research assessment, named *Valutazione della Qualità della Ricerca* (VQR).<sup>2</sup> The ASN was then taken in charge by the Ministry of Education and University, while ANVUR has steered two editions of the VQR, assessing research produced through 2004-2010 and through 2011-2014, respectively. Motivated by the explicit goal of compensating for late implementation, governments are making abundant use of the results of such procedures. In particular, the results of VQR have been employed to regulate a large share of public funding for universities, as well as for other issues that were beyond its original scopes (e.g. to set the standards for allowing professors to teach at some PhD-level courses).

The introduction of those formal evaluation systems by ANVUR raised several harsh criticisms. Most of them focused either on the ideological roots of evaluation (most notably Pinto 2012), or on technical mistakes

<sup>1</sup> See <http://abilitazione.miur.it/public/index.php?lang=eng>, accessed 12/12/2017.

<sup>2</sup> See [http://www.anvur.org/index.php?option=com\\_content&view=article&id=1206:vqr&catid=2:non-categorizzato&lang=it&Itemid=789](http://www.anvur.org/index.php?option=com_content&view=article&id=1206:vqr&catid=2:non-categorizzato&lang=it&Itemid=789), accessed 12/12/2017.

made by ANVUR (e.g. Banfi and De Nicolao 2013; Baccini 2016). However, in this paper we want to highlight another problematic feature of formal evaluation systems such as the Italian one. Namely, the risk that they could impoverish the epistemic pluralism of scientific communities. Some reflections were devoted to this issue by one of the very people who designed these formal evaluation, a former member of the ANVUR Steering Committee, in a book conceived as a rebuttal of several criticisms of formal evaluation (Bonaccorsi 2015; see also Bonaccorsi 2018).

Recognizing that some scientific communities, particularly in social sciences and humanities, may not share uniform methodological standards, Bonaccorsi acknowledges the possibility that evaluation may generate epistemic conflicts between different schools of thought that co-exist within a discipline. Advised by the philosopher Carla Bagnoli, he acknowledged the possibility that social mechanisms for value attribution may produce what Miranda Fricker (2007) called *epistemic injustice* (Bonaccorsi 2015, 76, fn. 11).

Notably, in order to introduce the concept of testimonial injustice,<sup>3</sup> Miranda Fricker presented the following example, inspired by a discussion with a scientist friend:

Imagine [...] a panel of referees on a science journal who have a dogmatic prejudice against a certain research method. It might reasonably be complained by a would-be contributor that authors who present hypotheses on the basis of the disfavoured method receive a prejudicially reduced level of credibility from the panel. Thus, the prejudice is such as to generate a genuine testimonial injustice. (Fricker 2007, 27)

However, while Bonaccorsi (2018) explicitly acknowledges and even endorses epistemic pluralism (at least in social sciences and humanities), he is optimistic that such injustices can be avoided by finding a common ground for assessing the research quality across different schools of thought. Contrary to his optimism, in this article I claim that epistemic pluralism is likely to be compromised by a bias that is rooted in any kind of evaluation – but gets amplified when evaluation procedures are highly formalized.

The discussion will proceed as follows: first, I will motivate the claim that epistemic pluralism is a desirable feature for the social organization of science. Then, I will describe a possible kind of epistemic bias that may negatively affect epistemic pluralism and show that it is very likely

---

<sup>3</sup> Fricker describes two varieties of epistemic injustices: testimonial injustice, occurring when a speaker is given less epistemic authority than she deserves; and hermeneutical injustice, occurring when a social group lacks the conceptual resources to make sense and to express its social experience.

to be at work in all sorts of evaluative practices. Therefore, since even informal evaluation might endanger epistemic pluralism, *a fortiori* particular care should be made with formal evaluation. I conclude by briefly hinting at some strategy that might be adopted to counter this risk.

### 3. Epistemic Pluralism

#### 3.1. What is Epistemic Pluralism

In science, the word ‘pluralism’ may refer to many different, if partially overlapping, concepts. For instance, we can have the following three kinds of pluralist stance:

- a) *Ontological pluralism*. Contrary to the neo-positivist ideal of unifying science by reducing special sciences to physics (Oppenheim and Putnam 1958), post-positivist philosophy of science argued in favour of a plurality of unreducible ontologies (e.g. Fodor 1974; Suppes 1978; Dupré 1993).
- b) *Sociological pluralism*. Feminist philosophy of science has denounced the underrepresentation of some social groups in science (e.g. women and ethnic minorities), and argued in favour of a more balanced composition, for both epistemological and political reasons (Anderson 2015).
- c) *Disciplinary pluralism*. More recently, some researchers compared the scientific productivity of various countries, revealing that it was higher in those that diversify their research efforts across more domains as opposed to specialising in some specific one; therefore, they suggest that national science policy-making should try to promote a pluralism of domain of inquiries (Cimini, Gabrielli, and Sylos Labini 2014).

Notwithstanding their relevance, the abovementioned kinds of pluralistic stances are not addressed in the present discussion. Rather, I focus on a fourth variety of pluralism, which I dub *epistemic pluralism*. My working definition is the following:

epistemic pluralism = the compresence of two or more rival schools of thought within a same domain of inquiry.

Given the lack of undisputable criteria for setting the boundaries between ‘rival schools of thoughts’, I shall settle for the following stipulation:

rival schools of thought = distinct research groups who endorse conflicting conceptual and/or methodological commitments, but whose explanatory scopes are at least partially overlapping – i.e., they are competing to explain some shared set of phenomena.

Usually – though not necessarily – such rivalries are revealed by different institutional features such as distinct scientific societies or scientific journals for each competing schools of thought, or by pragmatic features such as different technical languages (or if you prefer, ontologies) and heuristics. Their peculiar disagreement is not much about what they hold to be true about the world – members of a same school of thought may also disagree on that – but rather about *how to verify* these truths, i.e. by means of which methods, heuristics, models, idealizations. In a nutshell, what is at stake here is not the disagreement between specific theories *per se*, but rather between second-order conceptual frameworks – be them construed as thought collectives (Fleck 1935), paradigms (Kuhn 1962/1970), research programs (Lakatos 1970), research traditions (Laudan 1977) or something else. To name but a few intuitive examples of actual rivalry, think about psychodynamics and cognitive psychology; continental and analytic philosophy; neoclassical and heterodox economics.

Is epistemic pluralism desirable for science? If so, in which form? Divergent opinions existed in classical 20<sup>th</sup> century epistemology. Notably, describing the convergence on a single paradigm as a prerequisite for normal science, Thomas Kuhn (1962/1970) interprets the co-existence of rival schools of thought within a same discipline as a cue of immature science. However, while his justification for such convenience may be regarded as a transcendental argument for endorsing epistemic monism on a synchronic plane, his praise of ‘progress through revolution’ qualifies him as a supporter of diachronic pluralism (Viola 2015). Other philosophers held that epistemic pluralism should be pursued also on a synchronic plane. Notoriously, Paul Feyerabend (1975) argued for a very radical form of pluralism, expressed in the slogan ‘anything goes’ (see also Kellert, Longino and Waters 2006). However, one needs not commit to such radical stances to defend epistemic pluralism. More modestly, siding with Lakatos, one can recognize that

[t]he history of science has been and should be a history of competing research programmes (or, if you wish, ‘paradigms’), but it has not been and must not become a succession of periods of normal science: the sooner the competition starts, the better for progress. ‘Theoretical pluralism’ is better than ‘theoretical monism’: on this point Popper and Feyerabend are right and Kuhn is wrong (Lakatos 1970, 60).

Epistemic pluralism does not entail antirealism, nor ontological irreducible pluralism such as Duprè’s (1993): while being a realist, one could still maintain some form of convergent realism (Kellert, Longino and Waters 2006), holding that while in the long run the one true ontology will be discovered, no option should be foreclosed in advance.<sup>4</sup>

---

<sup>4</sup> I thank Giuliano Torrenco for pointing this out to me.

In the following sub-sections, I summarise some discussions concerning the desirability of pluralism in social epistemology. Analytical and agent-based models both seem to speak in favour of epistemic pluralism. Nonetheless, since the interpretation of these models is by no means straightforward, I will turn to another source for a modest defence of pluralism – namely, theoretical arguments based on historical considerations.

### 3.2. Why Epistemic Pluralism matters #1: economic epistemology

Starting from the nineties, social epistemologists began to reflect both on the desirability of epistemic pluralism and on the conditions that may promote or reduce it. An exemplar case of Goldman's (2011) system-oriented social epistemology, this literature addresses both normative and descriptive concerns. Moreover, it borrows some methods from social science – in particular, mathematical and agent-based modelling from economics: that is why it gained the label of 'economic epistemology' or 'economics of scientific knowledge' (Zamorra-Bonilla 2012).

Concerning the present discussion, the first milestone was posed by Kitcher in his 1990 article *The Division of Cognitive Labor*. Adopting a framework originally elaborated by Peirce (1879), Kitcher envisioned a scenario in which a community of scientists, who ought to make a given discovery, must choose if and how to split their cognitive efforts between two methods, i.e. they must answer the question: how many scientists should pursue each method?<sup>5</sup>

It is assumed that the probability that each method produces a discovery in a given timeframe depends on a return function that is increasing in the number of scientists. However, these return functions are concave: consequently, while a method  $M_1$  may be intrinsically superior to another method  $M_2$ <sup>6</sup>, overcrowding of the former can make it less efficient than hedging the bets. Therefore, hedging the bets by also having a minority of scientists who pursue  $M_2$  is wiser than having everybody pursuing  $M_1$ .

Kitcher then goes on discussing whether the social reward structure of science may be particularly fit for achieving this optimal allocation. Strevens (2003) further expands that point by comparing alternative reward structures, and claiming that the one that is more likely to sustain

---

<sup>5</sup> Kitcher's discussion is based built on the example of the discovery of the molecular structure of some molecule, which can be investigated either by empirical observation or by building toy-tinkers. While this example does not count as a genuine case of epistemic pluralism according to my definition, he specifies that his discussion is meant to refer to various kind of 'cognitive objects' such as "set of rival theories, research programs, methods for approaching a problem, etc." (Kitcher 1990, 10).

<sup>6</sup>  $M_1$  is superior to  $M_2$  if, given any number of scientists pursuing only one method, the probability of the discovery is always superior when this method is  $M_1$ .

the optimal allocation of scientists is indeed the actual system, based upon the *priority rule* (first described by Merton 1957), which prescribes that only the first one(s) who make a discovery get a prize for it ('winner takes all'). According to Strevens, this is the most rational allocation, because nothing would be gained by making the same discovery twice (but see §2.4). Nonetheless, Strevens in a later paper (2013) acknowledges that this optimistic assessment of the division of cognitive labour 'naturally' emerging from the adoption of the priority rule may be seriously endangered by the presence of herding behaviour. He notices that the 'golden share' for undertaking the correct scientific project often takes a (indeterminately) long time to unfold. Therefore, risk aversion might drive scientists to settle for more modest sources of credit, such as the recognition of their peers – typically expressed in the form of citation. But then, being into a more crowded school of thought make it easier to be recognized by a wider number of peers – thus making mainstream schools of thought more appealing than it would be rationally desirable.

Muldoon and Weisberg (2011) refer to these mathematical models as a Marginal Contribution/Reward approach, and criticize them for relying on controversial assumption (mostly inherited from classical economics: see Hands 1997; Mäki 2005; Viola 2015; Fernández-Pinto 2016).

### 3.3. Why Epistemic Pluralism matters #2: agent-based models

Instead of these models, Weisberg and Muldoon propose to investigate the division of cognitive labour through an agent-based model where the agents (scientists) must explore a three-dimensional 'epistemic landscape', representing the many possible approaches within a scientific field (Weisberg and Muldoon 2009). The 'landscape' is composed by many patches, each one representing a different approach within a given domain of inquiry. Some of these patches are higher than others (representing more fertile approaches), delineating some 'hills' of scientific fruitfulness. Agents ought to explore as many patches as they can among those whose epistemic significance above 0; to put it bluntly, they must climb the hills and its epistemically significant surroundings as soon as possible.

Each agent has only limited information: it only knows the epistemic significance of the patch it occupies, as well as that of those adjacent patterns that have been already explored by some other agent (i.e., scientists left traces of the in form of publications about the patterns they explored). However, Weisberg and Muldoon designed two kinds of agent<sup>7</sup>, distinguished by different behavioural patterns: *followers*, who

---

<sup>7</sup> Plus, a third kind used as control, that I ignore here.

follow the trails of other agents; and *mavericks*, who privilege the exploration of yet unknown patterns over the known ones.

Populations made entirely by followers perform worse than those made entirely by mavericks, because followers tend to cluster and get stuck in low significance regions instead of making brave explorations as mavericks do.<sup>8</sup> However, mixed populations perform even better.

Being aware of the high level of abstraction of their model, Weisberg and Muldoon refrain from drawing strong lessons out of it. Nonetheless, further considering that being a maverick is costlier than being a follower, they propose the tentative conclusion that “optimum research communities are going to be composed of a healthy number of followers with a small number of mavericks” (251). For the sake of the current debate, their conclusion can be interpreted as an indirect endorsement of epistemic pluralism (represented by the exploration of several patches), paired with the suggestion that pluralism is easier to achieve when scientists are biased toward the exploration of unknown approaches.

Other intriguing agent-based simulations with different architectures have been proposed, that are either moderately in favour or against synchronic epistemic pluralism. In Baliatti, Mäs and Helbing’s (2015) model, scientists ought to find a scientific truth, represented as a point within a bi-dimensional space. Scientists are ‘dragged’ along the space by the combined effect of three vectoral forces: first, they are attracted by the intrinsic force of the truth-point; second, they are influenced by their neighbour colleagues directions, which they mimic, provided that these colleagues stand within a given agent’s ‘sensory range’; third, for each agent there is some noise, i.e. some random force. According to this model, a marked epistemic pluralism – represented by scientists being sparsely distributed all over the landscape – hampers progress toward the truth because of the lack of forces that are strong enough to prevent self-reinforcing herding behaviour. In fact, in this scenario noise might deviate small clusters of researchers toward the wrong direction, arguably representing the self-reinforcement of prejudices held by a sub-community due to mimicking behaviours.

Instead, Zollman (2010) argues in favour of transient diversity. He models a scientific community as a network of interconnected Bayesians who play two-armed bandits. Each arm represents a different scientific approach and is characterized by a different payoff distribution. The payoff structure, however, is unknown to scientists. Rather, they have prior beliefs about which one is the better arm and update them by considering both the result of their own choices and those of the scientists

---

<sup>8</sup> In a follow-up, Weisberg (2013) refers to such a phenomenon as of *herding behavior*. He shows that while some strategies may reduce the herding behavior of followers, none of them would make them as efficient as mavericks.



they are linked with in the network. After testing several kinds of networks, Zollman concludes that while networks with *less* connections are slower, they are more reliable in making everyone converge on the (objectively) better arm, whereas highly interconnected networks sometimes converge on a self-reinforcing consensus over the wrong answer. However, a community with stubborn scientists having extreme priors will manage to test alternative hypotheses without discarding them too soon, and eventually it will converge on the right outcome even in highly interconnected networks. Yet, combining extreme priors and low interconnections tend to fossilize the disagreement and paralyze various clusters of scientists into their prejudice, thus failing to achieve consensus.

All things considered, while both Zollman and Weisberg and Muldoon's models suggest that a certain amount of epistemic pluralism might be beneficial (at least in some conditions), *prima facie* Bailetti and colleagues' model seems to point toward the opposite conclusion. This disagreement mainly depends on the different scopes and assumptions made by these models. Given that these model assumptions are presently "still rather disconnected from the real-world social organization of scientific research" (Martini and Fernández-Pinto 2016), we would refrain from drawing strong conclusion from them.

### **3.4. Why Epistemic Pluralism matters #3: historical and sociological considerations**

Given the uncertainty surrounding the models found in social epistemology, my endorsement of epistemic pluralism will mainly bear on two more modest epistemological arguments inspired by some simple historical and sociological considerations. I dub them the *prudence argument* and the *convergence argument* and discuss them in turn.

[*Prudence*] we cannot reliably foresee which one, among many rival schools of thought, is more likely to produce correct or more significant findings.

Often, a school of thought may fail to explain some phenomena which are easily accounted by another one. In contrast to what is assumed in Kitcher's (1990) model, the history of science seems to suggest that we cannot reliably foresee which school of thought is more likely to produce a given answer. This is vividly expressed in the case discussed by Zollman (2010), i.e. the discovery that peptic ulcers are caused by the *helicobacter pylori*. In 19<sup>th</sup> century, two competing hypotheses were proposed to explain the disease: the presence of some unobservable bacteria and an excess of acid. When in 1954 the prominent gastroenterologist Palmer published a study that appeared to demonstrate

that no bacteria can colonize the human stomach, this was taken as a conclusive evidence against the bacterial hypothesis. Sadly, his conclusion was unwarranted, since the kind of stain he used to investigate the biopsies was ‘blind’ to the *H. pylori*. It took about thirty years to Marshall and Warren to discover that the disease was caused by a bacterium. Yet, at first their discovery (yielding them the Nobel for Medicine in 2005) was dismissed by the medical community, since Palmer’s conclusions had been crystalized into received wisdom among the medical community. Fortunately, the frustrated Marshall behaved as Zollman’s stubborn scientist, and he himself drank a solution containing *H. pylori*. He manifested the symptoms of the peptic ulcer, and then effectively cured himself with an antibiotic, thus convincing his peers. Despite this story has a happy ending, Zollman cannot help wondering about how many more patients could have been successfully cured if only the bacterial hypothesis was not dismissed too soon. And we may also ask: how many correct hypotheses could have been overlooked if they have had no stubborn advocates such as Marshall?

Nonetheless, science must not only find truths: indeed, it should find *significant truths* (Kitcher 1993, 94). However, the significance of some scientific discoveries (which I take to indicate their potential to contribute to social well-being) cannot be unequivocally estimated *ex ante*, also because each piece of the puzzle of science might gain value depending on the availability of other pieces. This dynamic character of epistemic significance is nicely explored by Avin (2015a, ch. 4).<sup>9</sup> For instance, Avin stresses how the laser gyroscope, which required advancements in engineering and in theoretical physics made in the Sixties in order to be built, was only conceivable because of an experiment made in 1913 by Georges Sagnac, and published in France, whose original scope was to test *ether wind*.

Furthermore, many important discoveries in science were not due to some specific theory-testing. Rather, many ground-breaking discoveries emerged as the unexpected result of some fortuitous event – a circumstance for which the word *serendipity* has been coined. An evocative example is the discovery of penicillin: Alexander Fleming noticed that the cover of a Petri dish containing bacterial culture had not been properly set, and that a mould had grown, killing the bacteria.

Arfini, Bertolotti and Magnani stress that in order to make a serendipitous discovery, it needs to be “not expected, but [...] still recognizable, at least to certain cognitive systems. Otherwise, it would be pushed aside by consciousness”. Since the school of thoughts scaffold scientists’ cognitive system, a monopolistic school of thought with an overly restrictive ontology might work as a blindfold for some phenomena that are not predicted by its ontology. Indeed, according to Kuhn (1962), this is the

---

<sup>9</sup> This facet of Avin’s work was brought to my attention by Carlo Debernardi.

routinely way to deal with anomalies. Kyle Stanford (2015) makes a similar point, expressing the worry that the actual structure of science, due to the concentration of incentives toward conservative research, might make it harder to grasp unconceived alternatives.

These considerations stress the risks of allowing for the monopoly of a single school of thought, thus vindicating a cautionary rationale for preserving at least a *minimal epistemic pluralism*. Lastly, it is worth keeping in mind that scientific activity is imbued with *tacit knowledge* (Polanyi 1966). This kind of knowledge is hardly translatable onto explicit knowledge; rather, it is usually transmitted through long apprenticeship – which is why reading textbooks is not enough to become a scientist, but a doctorate or some other equivalent form of apprenticeship is in order. Thus, allowing a school of thought to completely extinguish likely implies a loss of tacit knowledge – perhaps right before the availability of some piece of the puzzle would make it priceless: a despicable loss, compared to the relative small price of letting some minority school of thought continue its legacy, if just for a few stubborn followers.

[*Convergence*] if a multiplicity of independent rival schools of thought converges on a same result, this result is more reliable.

Strevens's (2003) abovementioned defence of the priority rule was based upon the assumption that we do not need the same discover to be made twice. However, since the reliability of science significantly bears on the reproducibility of its findings,<sup>10</sup> giving no incentives at all for replications is tantamount to deprive science of its antibodies, because scientific frauds and mistakes will lurk for longer, and perhaps forever – a topic which is daunting for current research (e.g. Ioannidis 2005).

In particular, the better guarantee for the reliability of (a piece of) scientific knowledge comes from the convergence of many independent sources (Kosso 1989). Jean Perrin's discovery of the Avogadro's number counts as an exemplar case of a reliable knowledge, since he “measured the same physical quantity in a variety of different ways, thereby invoking a variety of different auxiliary theories” (Kosso 1989, 247). Given that rival schools of thought employ, by definition (see above), different methods for testing scientific statements, it follows that whatever scientific theory is deemed true by distinct, even rival schools of thought, is *ceteris paribus* more robustly validated than one that is backed solely by the followers of a single school of thought.

---

<sup>10</sup> Eugenio Petrovich brought to my attention that now that we are in the era of Big Science, where some experiments are simply too expensive to be replicated (think about the CERN in Genève), replication might belong to the idealized image of science, rather than to its accurate description (Collins 1992). I concede that this might be true when it comes to big science. Yet, I contend that replication remains possible in many domains outside Big Science – and that failures to replicate present serious reasons of concern.

## 4. Epistemic Bias

### 4.1. Why should evaluations be threatening to Epistemic Pluralism?

Evaluation is usually aimed at measuring the scientific quality of some research products and/or of those who produced it. However, provided that scientific quality exists, what makes us think that evaluators would be able to grasp it objectively, without being affected by their prejudices? Even excluding the case of deliberate boycott of rival schools of thought by scholars engaged in evaluation procedures, it is well established that human beings are prone to many implicit biases. Why then should we suppose that scientists are immune to biases when assessing their peers and their work? And what happens if such biases are embedded into large-scale formal evaluation with profound implications for a national scientific environment?<sup>11</sup>

Some reflections on biases that may harm epistemic pluralism can be found in Donald Gillies's 2008 book. The book is a critical assessment of the Research Assessment Exercise (RAE, recently replaced by Research Excellence Framework, REF), whose results have been used to allocate public funding for research in the UK. Since RAE was based on peer review, Gillies asks: is peer review able to predict which research projects are going to bear fruitful results? Were his answer positive, it could be reasonable to concentrate many resources to that project, even at the expense of other strands. But that seems not to be the case. As he summarized in a later article,

the root of the problem is what I will call *researcher narcissism*. This is a condition, which affects nearly all researchers (including the author of the present paper). It consists in an individual researcher believing quite strongly that his or her approach to research in the field is the best one, and most likely to produce good results, while the other approaches are less good and less likely to produce any good results (Gillies 2014, 8).

Gillies adopts a counterfactual strategy to substantiate his scepticism: he discusses various historical cases from many fields where peer review would have failed to foresee ground-breaking scientific advances: Frege's invention of mathematical logic, Copernicus's theorisation of heliocentrism, Semmelweis's invention of antisepsis (Gillies 2008, ch. 3), James Black's Nobel-awarding inventions of two important drugs (Gillies 2014).

---

<sup>11</sup> Feminist philosopher Jennifer Saul (2012) addressed the issue of implicit biases in two evaluative exercise undergone in the UK: the Philosophical Gourmet Report and the Research Excellence Framework. However, since she is only concerned with what I called sociological pluralism (§1), I do not address her work here.

## 4.2. Is there evidence for Epistemic Bias?

Despite its intuitive appeal, Gillies's discussion on researcher narcissism is purely speculative. Is there any evidence that such a kind of *epistemic bias* is at play in evaluation? Is it stronger in formal evaluation? To address these questions, I browsed the literature in several social sciences that deal with the presence of biases across three different kinds of evaluative practice: peer review, bibliometric evaluations and hiring procedures.

### *Peer review*

Peer review has been compared to democracy; both have been described as “a system full of problems but the least worst we have” (Smith 2006). Among these problems, many researchers highlighted many biases that compromise the alleged impartiality of the process (see Langfeldt 2006; Lee, Sugimoto, Zhang, and Cronin 2013). Nonetheless, few scholars specifically addressed the issue of epistemic bias as distinct from other biases, also because of the difficulty to disentangle them.<sup>12</sup>

If such epistemic bias applies to peer reviewing of scientific articles, referees might simply reject papers from rival schools of thought they disagree with, or even steer the author toward their own theoretical perspective. Might this be the case?

Some evidence in the literature suggests that the answer might be ‘yes’. Mahoney (1977) asked 75 (unaware) referees in experimental psychology to review a given manuscript, whose results he slightly modified, along with their interpretation. He found that referees tend to judge more positively the manuscripts that show positive results and/or that are in line with the theoretical perspectives of the referees.

However, whereas epistemic bias can exert a significant effect on scientific careers by influencing the fate of articles submitted in prestigious journals, its role is even more direct when it comes to allocate research fund. It is difficult to disagree that funding agencies may affect “the cognitive development of science by the structuring of the way in which research is done” (Braun 1998, 810; see also Goldman 1999, 257). Sadly (for epistemic pluralism), in doing that, they foster conservative researches over innovative ones (Braun 1998; see also Berezni 1998).

Having attended some meetings of panels that ought to adjudicate grants on behalf of the *Science and Engineering Research Council* (SERC), Travis and Collins (1991) observed that “committee members sometimes

---

<sup>12</sup> Wang & Sandström (2015) tried to overcome this problem by developing a data mining technique to measure the ‘cognitive distance’ between researchers, comparing how much both their references and some key term from their abstract overlap. However, their measurement conflates the proximity of school of thought with that of topic.

make decisions based upon their membership in scientific schools of thought” (323).

More recently, Luukkonen (2012) wondered whether ERC panels were able to ensure that funds are channelled into “new and promising areas of research with more flexibility” (<http://erc.europa.eu/mission>). Her answer was negative: she declared that “despite the ERC’s aims, the peer review process in some ways constrains the promotion of truly innovative research” (Luukkonen 2012, 11), and she further observes that “[t]hese constraints arise from the very essence of peer review, namely, its basic function of judging the value of proposed research against current knowledge boundaries” (*ibid.*).

### *Bibliometrics*

*Prima facie*, due to their mathematical format, bibliometric indicators might seem good candidates for providing *objective* measures of research quality. They also might be tempting due to their relative inexpensiveness, especially in large scale formal evaluation exercises where the number of research products to evaluate is high. However, it is worth remembering that since most widespread bibliometric indicators (e.g. *impact factor* and *h-index*, respectively measuring the impact of journals and of researchers) are based upon counting citations within peer reviewed articles indexed by a given database, they embed and aggregate the prejudices of both the referees and the editors of the journals, plus the indexing criteria of the database owners. Moreover, bibliometric indicators are meant to represent *impact*, not *quality*. Whereas sometimes impact is considered as a reliable proxy of quality (e.g. in the Italian VQR it is done for many scientific fields, especially in hard and life sciences), this identification is problematic, as it provides strong disincentives to work on mainstream problems and within heterodox schools of thought (as documented for instance by Castellani, Pontecorvo, and Valente 2016). To understand why this happens, consider the following scenario: two papers of comparable quality, P1 and P2, provide a relevant insight over a same issue. However, P1 does so from the standpoint<sup>13</sup> of a mainstream school of thought, with huge number of followers, whereas P2 from that of a minor (or yet to exist) school of thought, with far less followers. Given the reasonable assumption that the segregation between school of thoughts make it relatively less likely that some scholar reads (and thus cite) a paper from a rival school of thought, all else being equal, the wider audience would boost P1 impact far over P2’s – irrespectively of their quality. In Muldoon and Weisberg’s (2009)

---

<sup>13</sup> Here the notion of ‘standpoint’ should be interpreted according to either a cognitive dimension (e.g. the lexicon used to frame a problem, the method used to test it) or a sociological one (e.g. the paper being discussed in some conferences and published in journals that are commonly associated with one school of thought).

terms, the ‘citation economy’ discourages people from behaving like mavericks, because mavericks are arguably less likely to get cited.

For these and other reasons, many institutions and scientists subscribed the San Francisco Declaration on Research Assessment, which prescribe “not [to] use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist’s contributions, or in hiring, promotion, or funding decisions” (<http://www.ascb.org/dora/>).

### *Hiring*

Among the various evaluative procedures, comparative evaluations of candidates for academic recruitment, as well as procedures assessing their eligibility (e.g. the abovementioned ASN in Italy), are possibly the more relevant as for epistemic bias due to a path-dependent reinforcing loops. In fact, it is very likely that the promoted candidate will oversee judging some future candidates, and any epistemic bias will be transmitted to the next generation of evaluators – and thus amplified.

Available literature shows that some biases are indeed at play during hiring procedures: for instance, candidates who are somehow connected with the examiners (e.g. they are co-authors of some articles, or come from a same department), are more likely to be hired (e.g. for France see Combes, Linnemer, and Visser 2008 and Godechot 2016; for Spain, see Zynovieva and Bagues 2015). However, these authors stress that they are unable to judge whether this advantage was due to epistemic bias or rather to social particularism (such as nepotism).<sup>14</sup>

Nonetheless, even social biases may have important epistemic consequences. Studying the hiring networks of American research institutions in computer sciences, business, and history, Clauzet, Arbesman, and Larremore (2015) highlighted a very ‘endogamous’ and hierarchical structure: on the one hand, most professors (about four out of five) obtained their PhD in one of the ‘top’ 25% departments – among which further hierarchical layers could be distinguished. On the other hand, almost none of those who obtained their PhD in less prestigious institutions managed to be hired at the higher-levels. Therefore, they conclude that

the centralized and highly connected positions of higher-prestige institutions enable substantial influence, via doctoral placement, over the research agendas, research communities, and departmental norms throughout a discipline . . . . The close proximity of the core to the entire network implies that ideas originating in the high-prestige core, regardless of their merit, spread more easily through-

---

<sup>14</sup> I speculate that in future works this confound may be at least partially addressed by applying formalized measures of cognitive distance such as that developed by Wang and Sandström (2015; see fn. 12).

-out the discipline, whereas ideas originating from low-prestige institutions must filter through many more intermediaries (Clauzet et al. 2015, 4).

To sum up, there is moderate evidence that epistemic bias is at play within each of the three evaluative activities I have discussed, i.e. peer review, bibliometrics, and hiring procedures. A sceptic may still argue that this evidence is far from being conclusive. I concede that. However, since I think that most researchers would take for granted that epistemic bias is at play and significantly distorts evaluations, I claim that the burden of disproof is up to the sceptics. Moreover, even if the pluralism-reducing effect of epistemic bias were moderate in each of the above-mentioned fields, the cumulative impact may be significant: even though being hired were only *slightly* more difficult if the members of the panel are hostile to your school of thought, it does become *considerably* harder if due to the unpopularity of your school of thought you had an hard time publishing your papers in high-ranked journals and to get them cited. Such a self-reinforcing loop has been reported to affect economics schools of thought in the UK: according to Lee, Pham, and Gu (2013), twenty years of RAE resulted in the disappearance of heterodox rival schools of thought in favour of mainstream economics.

All things considered, unless and until sceptics succeed in demonstrating that epistemic bias is negligible or harmless, there are strong reasons for worrying about epistemic bias and for trying to mitigate it.

## 5. Formal evaluation enhances epistemic bias

In §3 I have defined epistemic pluralism and argued in favour of at least a minimal form of pluralism. Then, in §4 I have introduced and substantiated the hypothesis that any kind of evaluative process is prone to be affected by epistemic bias, i.e. the evaluators might favour those who pursue their same school of thought over those who do not.

The simplest and most radical solution would be to cease any evaluation. Yet, this is hardly a viable option: as far as some finite amount of public resources must be allocated, we need some criteria for choosing how to allocate them. However, while *some* form of evaluation is mandatory, these forms need not be strong evaluation systems that (a) are steered by some rather restricted scientific elite, (b) follow highly formalized rules and procedures, and (c) have a straightforward impact on affect funding and careers. These are the characteristics of the systems described by Hicks (2012), and especially of the Italian systems described in §2. Due to their often wide-scope, they make a large use of bibliometry (e.g., in Italy bibliometric index have been employed for hard and life sciences) or to other highly standardised index and rankings (e.g., in Italy journals in social and human sciences have been classified in hierarchical rankings



for the ASN). According to Whitley (2007, 10), such systems tend to impose a standardization and a institutionalization of goals and values to a scientific discipline, so that “the diversity of intellectual goals and approaches within sciences should decline over time, especially where they challenge current orthodoxies”. Eventually, “such reinforcement of disciplinary standards and objectives is likely to inhibit the development of new fields and goals that transcend current intellectual and organisational boundaries by increasing the risks of investing in research projects that do not fit within them” (*ibid.*).

My hypothesis is that this happens because formal evaluation amplifies the epistemic bias already existing in weaker evaluation practices, as well as accelerating their pluralism-dampening effects. This is consistent with the claim of Bonaccorsi (who recently governed the implementation of such procedures in Italy) that formal evaluation systems “ha[ve] the effect to foster and catalyze the epistemic reflection of the community” (2015, 88, translation is mine). However, Bonaccorsi does not side against epistemic pluralism, that he recognizes as a genuine (and perhaps even beneficial) feature of social sciences and humanities. He is optimistic that epistemic pluralism might be preserved notwithstanding epistemic bias, because he deems possible that schools of thought find some common ground for bias-free evaluation. Though, on the light of the evidence of epistemic bias discussed in §4, it seems much more likely that a dominant school of thought will impose its evaluative criteria as a common ground, promoting the extinction of scientific minorities (or preventing the birth of new ones). This evidence is not conclusive, but is likely sufficiently strong to put the burden of proof upon the shoulders of those who deem, like Bonaccorsi, that epistemic bias can be made consistent with a common ground for evaluation.

## **6. Some hypotheses for protecting epistemic pluralism from epistemic bias**

Possibly, epistemic bias cannot be completely counteracted. However, some strategies have been proposed in order to reduce it. Bonaccorsi (2015) concedes that if (and only if) a common ground cannot be found (and there are reasons to suspect that this will be the norm, rather than the exception), members of the evaluative panels must be selected with the aim to represent (m)any school(s) of thought. He also stresses the importance of a rapid turnover of these panels.

Drwaing upon research in management studies, Osterloh and Frey (2015) endorse a more radical answer to the question “what kind of control is suited for science?” They think that both output control, i.e. bibliometrics, and process control, i.e. peer review, have too many flaws, and produce too many distortions. The only opinion left is input control:

in their opinion, candidate researchers should undergo a thorough hiring procedure, but then, if they get hired, they should be left free to determine their agenda by themselves (cf. Gillies 2008).

As for research grants, it has been proposed to supersede epistemic bias by picking at least some of them at random, through a lottery. This proposal has been recently detailed by Avin (2015a, 2015b, 2018; see also Gillies 2014), who also explained its rationale. To put it shortly: research project for grant allocation should be kept short, since long projects absorb plenty of time from both those who write and those who read them, and yet they fail to overcome the intrinsic unpredictability of the projects' outcome. All proposals of high merit should be funded, just as all proposals of low merit should be discarded. This, however, leaves out a wide number of proposals of medium merit. Given that noise and unpredictability would render finer-grained assessment useless, these medium-merit proposals should be placed in a lottery, and the winners should be funded. According to Avin, this method might lower the costs (especially of time), increase fairness and even make unorthodox ideas more easily funded.<sup>15</sup>

Other thinkers have recommended to fund people, rather than projects (e.g. Berezin 1998) – a proposal that has been considered by many institutions of the NIH in the US (Kaiser 2014). To begin with, it could be wise not to concentrate all the funding into a single agency (Travis and Collins 1991): as reported by Whitley (2007), diversification of funding sources might soften the effects of formal evaluation systems.

Be as it may, the arguably most efficient strategy for slowing down the effects of epistemic bias, thereby preserving epistemic pluralism, is that of inverting the actual trend of concentrating resources in the hands of few researchers at the expenses of the many (Sylos Labini 2016). This might also be achieved by mitigating the use of formal evaluation in allocating funds, or simply by doing without them altogether.<sup>16</sup>

The assessment of the merits and flaws of these and other proposals would require a thoroughly discussion based on empirical analyses that also takes into account contextual information. Obviously, such an endeavour lies beyond the purpose of the present article. Hopefully, such an assessment would take benefit from a careful and well informed public

---

<sup>15</sup> As reported by Avin (2018), some lotteries have been implemented already for some grants (the Health Research Council of New Zealand's "Explorer Grants", New Zealand's Science for Technology Innovation "Seed Projects" and the Volkswagen Foundation's "Experiment!" grants). However, their implementation is too recent to assess the efficacy of the policy.

<sup>16</sup> This might also lead to considerable savings: Geuna and Piolatto (2016) estimate that, all things considered, UK spent around 88.5 million € for RAE 2008, and are spending from 130 up to 164 million € for REF 2014, while Italy will spend about 70 million € for the second edition of VQR.

debate. Prompting it is the aim of this article.

## Acknowledgments

The author thanks the many people who provided insightful comments on earlier versions of this paper: Selene Arfini, Alberto Baccini, Raffaele Caterina, Carlo Debernardi, Giuseppe De Nicolao, Paola Galimberti, Alex Gillett, Eugenio Petrovich, Francesco Sylos Labini, Silvia Tossut, and the anonymous reviewers. The author is also grateful to Matteo Pinna Pintor for the proofreading, as well as for some interesting comments.

## REFERENCES

- Anderson, E. 2015. Feminist Epistemology and Philosophy of Science. *The Stanford Encyclopedia of Philosophy*. Edited by Edward N. Zalta. <http://plato.stanford.edu/entries/feminism-epistemology/>.
- Arfini, S., Bertolotti, T. and Magnani, L. 2018. *The Antinomies of Serendipity. How to Cognitively Frame Serendipity for Scientific Discoveries*, Topoi, DOI: 10.1007/s11245-018-9571-3
- Avin, S. 2015a. *Breaking the Grant Cycle: On the Rational Allocation of Public Resources to Scientific Research Projects* (Doctoral dissertation, University of Cambridge). Accessed December 15, 2017. [https://www.repository.cam.ac.uk/bitstream/handle/1810/247434/phd\\_dissertation\\_final\\_for\\_print.pdf?sequence=1&isAllowed=y](https://www.repository.cam.ac.uk/bitstream/handle/1810/247434/phd_dissertation_final_for_print.pdf?sequence=1&isAllowed=y).
- \_\_\_\_\_. 2015b. Funding Science by Lottery. In *Recent Developments in the Philosophy of Science: EPSA13 Helsinki*, eds. U. Mäki, I. Votsis, S. Ruphy and G. Schurz, 111-126. Basel: Springer International Publishing.
- \_\_\_\_\_. 2018. Policy Considerations for Random Allocation of Research Funds. *RT. A Journal on Research Policy and Evaluation* 6(1). <https://doi.org/10.13130/2282-5398/8626>.

- Baccini, A. 2016. Napoléon et l'évaluation Bibliométrique de la Recherche: Considérations sur la Réforme de l'Université et sur l'Action de l'Agence Nationale d'évaluation en Italie. *Canadian Journal of Information and Library Science* 40(1): 37-57.
- Balietti, S., Mäs, M. and Helbing, D. 2015. On Disciplinary Fragmentation and Scientific progress. *PloS one* 10(3): e0118747. <https://doi.org/10.1371/journal.pone.0118747>
- Banfi, A. and De Nicolao, G. 2013. Valutare Senza Sapere. Come Salvare la Valutazione della Ricerca in Italia da chi Pretende di Usarla Senza Conoscerla. *Aut Aut* 360: 43-68.
- Berezin, A. 1998. The Perils of Centralized Research Funding Systems. *Knowledge, Technology & Policy* 11(3): 5-26.
- Bonaccorsi, A. 2015. *La Valutazione Possibile. Teoria e Pratica nel Mondo della Ricerca*. Bologna: Il Mulino.
- \_\_\_\_\_ 2018. Towards an Epistemic Approach to Evaluation in SSH. *The Evaluation of Research in Social Sciences and Humanities*, ed. A. Bonaccorsi, 1-29. Cham: Springer.
- Braun, D. 1998. The Role of Funding Agencies in the Cognitive Development of Science. *Research policy* 27(8): 807-821.
- Castellani, T., Pontecorvo, E. and Valente, A. 2016. Epistemic Consequences of Bibliometrics-based Evaluation: Insights from the Scientific Community. *Social Epistemology* 30(4): 1-22.
- Cimini, G., Gabrielli, A. and Labini, F. S. 2014. The Scientific Competitiveness of Nations. *PloS one* 9(12), e113470. <https://doi.org/10.1371/journal.pone.0113470>
- Clauset, A., Arbesman, S. and Larremore, D. B. 2015. Systematic Inequality and Hierarchy in Faculty Hiring Networks. *Science Advances* 1(1), e1400005. <https://doi.org/10.1126/sciadv.1400005>
- Collins, H. 1992. *Changing Order: Replication and Induction in Scientific Practice*. Chicago: University of Chicago Press.
- Combes, P. P., Linnemer, L., and Visser, M. 2008. Publish or peer-rich? The Role of Skills and Networks in Hiring Economics Professors. *Labour Economics* 15(3): 423-441.
- Dupré, J. 1993. *The Disorder of Things. Metaphysical Foundations of the Disunity of Science*. Cambridge and London: Harvard Press.

- Fernández Pinto, M. 2016. Economics Imperialism in Social Epistemology: A Critical Assessment. *Philosophy of the Social Sciences* 46(5): 443-472.
- Feyerabend, P. 1975. *Against Method: Outline of an Anarchistic Theory of Knowledge*. Bristol: New Left Books.
- Fleck, L. 1979. *Genesis and Development of a Scientific Fact*. Chicago: University of Chicago Press. (Original: *Entstehung und Entwicklung einer wissenschaftlichen Tatsache: Einführung in die Lehre vom Denkstil und Denkkollektiv*, 1935)
- Fodor, J. A. 1974. Special Sciences (or: The Disunity of Science as a Working Hypothesis). *Synthese* 28(2): 97-115.
- Fricker, M. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. New York: Oxford University Press.
- Geuna, A. and Piolatto, M. 2016. Research Assessment in the UK and Italy: Costly and Difficult, but Probably Worth It (at Least for a While). *Research Policy* 45(1): 260-271.
- Gillies, D. 2008. *How should research be organised?* London: College Publications.
- \_\_\_\_\_. 2014. Selecting Applications for Funding: Why Random Choice is Better than Peer Review. *RT. A Journal on Research Policy and Evaluation* 2(1). <https://doi.org/10.13130/2282-5398/3834>
- Godechot, O. 2016. The Chance of Influence: A Natural Experiment on the Role of Social Capital in Faculty Recruitment. *Social Networks* 46: 60-75.
- Goldman, A. I. 2011. A Guide to Social Epistemology. *Social Epistemology: Essential Readings*, eds. A. I. Goldman and D. Whitcomb, 11-37. New York: Oxford University Press.
- Hands, D. W. 1997. Caveat Emptor: Economics and Contemporary Philosophy of Science. *Philosophy of Science* 64: S107-S116.
- Hicks, D. 2012. Performance-based University Research Funding Systems. *Research Policy* 41(2): 251-261.
- Kaiser, J. 2014. NIH Institute Considers Broad Shift to ‘People’ Awards. *Science* 345(6195): 366-367.
- Kellert, S. H., Longino, H. E. and Waters, C. K., eds. 2006. *Scientific Pluralism*. Minneapolis: University of Minnesota Press.

- Kickert, W. 1995. Steering at a Distance: A New Paradigm of Public Governance in Dutch Higher Education. *Governance* 8(1): 135-157.
- Kitcher, P. 1990. The Division of Cognitive Labor. *The Journal of Philosophy* 87(1): 5-22.
- \_\_\_\_\_. 1993. *The Advancement of Science: Science without Legend, Objectivity without Illusion*. Oxford: Oxford University Press.
- Kosso, P. 1989. Science and Objectivity. *The Journal of Philosophy* 86(5): 245-257.
- Kuhn, T. S. 1962/1970. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Ioannidis, J. P. 2005. Why Most Published Research Findings are False. *PLoS med* 2(8): e124. <https://doi.org/10.1371/journal.pmed.0020124>
- Lakatos, I. 1970. Falsification and the Methodology of Scientific Research Programmes. In *Criticism and the Growth of Knowledge*, eds. I. Lakatos and A. Musgrave, 8-101. Cambridge: Cambridge University Press.
- Langfeldt, L. 2006. The Policy Challenges of Peer Review: Managing Bias, Conflict of Interests and Interdisciplinary Assessments. *Research Evaluation* 15(1): 31-41.
- Laudan, L. 1977. *Progress and its Problems: Towards a Theory of Scientific Growth*. Berkley and Los Angeles: University of California Press.
- Lee, F. S., Pham, X. and Gu, G. 2013. The UK Research Assessment Exercise and the Narrowing of UK Economics. *Cambridge Journal of Economics* 37(4): 693-717.
- Lee, C. J., Sugimoto, C. R., Zhang, G., and Cronin, B. 2013. Bias in Peer Review. *Journal of the American Society for Information Science and Technology* 64(1): 2-17.
- Luukkonen, T. 2012. Conservatism and Risk-Taking in Peer Review: Emerging ERC Practices. *Research Evaluation* 21(1): 48-60.
- Mahoney, M. J. 1977. Publication Prejudices: An Experimental Study of Confirmatory Bias in the Peer Review System. *Cognitive Therapy and Research* 1(2): 161-175.

- Mäki, U. 2005. Economic Epistemology: Hopes and Horrors. *Episteme* 1(03): 211-222.
- Martini, C. and Pinto, M. F. 2017. *Modeling the Social Organization of Science*. *European Journal for Philosophy of Science* 7(2): 221-238.
- Merton, R. K. 1957. Priorities in Scientific Discovery: A Chapter in the Sociology of Science. *American Sociological Review* 22(6): 635-659.
- Muldoon, R. and Weisberg, M. 2011. Robustness and Idealization in Models of Cognitive Labor. *Synthese* 183(2): 161-174.
- Oppenheim, P. and Putnam, H. 1958. Unity of Science as a Working Hypothesis. In *Concepts, Theories and the Mind-Body Problem. Minnesota Studies in the Philosophy of Science. Vol. II.*, eds. H. Feigl, M. Scriven and G. Maxwell, 3-36. Minneapolis: University of Minnesota Press.
- Osterloh, M. and Frey, B. S. 2015. Ranking Games. *Evaluation Review* 39(1): 102-129.
- Peirce, Charles S. 1879. Note on the Theory of the Economy of Research. In *Report of the Superintendent of the United States Coast Survey Showing the Progress of the Work for the Fiscal Year Ending with June 1876*, 197-201. Washington DC: US Government Printing Office.
- Pinto, V. 2012. *Valutare e Punire. Una Critica alla Cultura della Valutazione*. Napoli: Cronopio.
- Polanyi, M. 1966. *The Tacit Dimension*. London: Routledge.
- Smith, R. 2006. Peer Review: A Flawed Process at the Heart of Science and Journals. *Journal of the Royal Society of Medicine* 99(4): 178-182.
- Stanford, P. K. 2015. Unconceived Alternatives and Conservatism in Science: The Impact of Professionalization, Peer-Review, and Big Science. *Synthese*. <https://doi.org/10.1007/s11229-015-0856-4>
- Strevens, M. 2003. The Role of the Priority rule in Science. *The Journal of Philosophy* 100(2): 55-79.
- \_\_\_\_\_ 2013. Herding and the Quest for Credit. *Journal of Economic Methodology* 20(1): 19-34.

- Suppes, P. 1978. The Plurality of Science. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, Vol. 1978*, 3-16. Chicago: University of Chicago Press.
- Sylos Labini, F. 2016. *Science and the Economic Crisis: Impact on Science, Lessons from Science*. Basel: Springer International Publishing.
- Travis, G. D. L. and Collins, H. M. 1991. New Light on Old Boys: Cognitive and Institutional Particularism in the Peer Review System. *Science, Technology & Human Values* 16(3): 322-341.
- Viola, M. 2015. Some Remarks on the Division of Cognitive Labor. *RT. A Journal on Research Policy and Evaluation*, 3(1). <https://doi.org/10.13130/2282-5398/4768>
- Wang, Q. and Sandström, U. 2015. Defining the Role of Cognitive Distance in the Peer Review Process with an Explorative Study of a Grant Scheme in Infection Biology. *Research Evaluation* 24(3): 271-281.
- Weisberg, M. 2013. Modeling Herding Behavior and its Risks. *Journal of Economic Methodology* 20(1): 6-18.
- Weisberg, M. and Muldoon, R. 2009. Epistemic Landscapes and the Division of Cognitive Labor\*. *Philosophy of Science* 76(2): 225-252.
- Whitley, R. 2007. Changing Governance of the Public Sciences: The Consequences of Establishing Research Evaluation Systems for Knowledge Production in Different Countries and Scientific Fields. In *The changing Governance of the Sciences: The Advent of Research Evaluation Systems*, eds. R. Whitley and J. Gläser, 3-30. Dordrecht: Springer Science & Business Media.
- Whitley, R. and Gläser, J., eds. 2007. *The Changing Governance of the Sciences: The Advent of Research Evaluation Systems*. Dordrecht: Springer Science & Business Media.
- Zamora Bonilla, Jesus P. 2012. The Economics of Scientific Knowledge. In *Handbook of the Philosophy of Science. The Philosophy of Economics*, ed. U. Mäki, 823-862. Amsterdam: Elsevier.
- Zollman, K. J. 2010. The Epistemic Benefit of Transient Diversity. *Erkenntnis* 72(1): 17-35.
- Zinovyeva, N. and Bagues, M. 2015. The Role of Connections in Academic Promotions. *American Economic Journal: Applied Economics* 7(2): 264-292.