

Chapter 11

Why Should Philosophers of Science Pay Attention to the Commercialization of Academic Science?

Gürol Irzik

11.1 Introduction

There is considerable evidence that since 1980 a new regime of science organization became dominant in the US, replacing the old one which was operative since 1945 (Etzkowitz and Webster 1995; Jasanoff 2005; Mirowski and Sent 2008). The old regime was formulated vividly in Vannevar Bush's famous 1945 report, *Science – The Endless Frontier*, according to which a simple division of labor between the state and the scientists was envisioned: while the former would set the research prerogatives and provide the funds, the latter would produce scientific discoveries which would then be developed into useful products by the industry for the benefit of the nation. In this mode of scientific knowledge production, universities would be the major actors in producing “basic” science and enjoy a high degree of internal autonomy and academic freedom.

Under the pressure of a number of forces, this old regime broke down. The new regime was established on the basis of an ever-expanding intellectual property rights, the privatization of publicly funded research, and new forms of collaboration between the university, the state and the industry (Bok 2003; Boyle 1997; Greenberg 2001; Krimsky 2004; Magnus et al. 2002; McSherry 2001; Mirowski and Sent 2008). It can be seen as responding to the demands of what is often called “post-industrial capitalism” or “knowledge economy”, to use a less politically charged phrase. The common assumption is that expert knowledge, which is above all scientific knowledge, has become a factor of production more important than labor, land and money and a key to economic competitiveness. As a result, scientific knowledge became commodified and certain segments of academic science, notably biomedicine and genetics, have become rapidly commercialized in unprecedented ways primarily in the US and to a lesser degree elsewhere.

G. Irzik (✉)
Sabanci University, Istanbul, Turkey
e-mail: irzik@sabanciuniv.edu

Although, in the last decade or so there has been an explosion of publications, drawing attention to and discussing various aspects of commercialization of academic science and commodification of scientific knowledge more specifically, not withstanding a few exceptions, philosophers of science have been largely impervious to this phenomenon. My main thesis is that this phenomenon has direct bearing on some of the central problems in the philosophy of science and that as philosophers of science we should be well advised to take its impact on science seriously.

The plan of my article is as follows. In Section 11.2 I outline briefly the political, economic, legal and scientific developments that led to this phenomenon in the U.S. Based on the existing literature on the topic, I argue in Section 11.3 that while commodification of scientific knowledge does make economies more competitive and productive, it also has a number of negative effects on certain aspects of science, such as the choice of scientific problems and the direction of scientific research, the social norms and the function of science. Commercialization also affects the distinctions between discovery and invention, between fact and artifact and between nature and culture, more broadly speaking. I discuss these in Section 11.4. I conclude with some general remarks.

11.2 Science for Sale: The Road to Commercialization

We are faced with a conceptual issue right at the beginning: how is this new regime of science organization to be understood and described? Several terms are used in the existing literature – “commodification of academic research”, “commercialization of (academic) science”, “globalized privatization regime” (see, for example, Radder forthcoming; [Mirowski and Sent 2008](#)). However we call it, it is an extremely complex and heterogeneous phenomenon that defies a simple definition, but the basic idea is that academic science begins to be commercialized when scientific research is done, scientific knowledge is produced, and scientific expertise is mobilized in the universities and other academic institutions primarily for the purpose of profit. When scientific knowledge is produced primarily for making money, we may speak of its commodification. Commodification of knowledge is made possible via intellectual property rights, in terms of patents, copyrights and licensing.

During most of the twentieth century, the prevalent attitude was that academic science and property did not go together. The latter was considered to be a notion antithetical to the scientific enterprise in the universities, and accordingly most university scientists were reluctant to patent the results of their inventions, especially when they concerned public health. As a result, many inventions were not patented. Two of the most important of these were magnetic resonance imaging and the polio vaccine. In line with this, most universities did not have any patent policies until after World War II and approached the issue of patents in health sciences unfavorably ([Irizik 2007](#)). Things began to change dramatically in the last three decades, however. While I cannot do justice the history of this complex change, I can summarize the factors behind it schematically as follows:

Economic-political: Since the 1970s the economies of the most developed countries entered a new phase and became “knowledge economies”, where expert knowledge became the major factor of production. At the same time, a global world market became a reality more than ever, and economic competition between countries at the global scale reached new heights. With Reaganism in the U.S. and Thatcherism in England, neoliberal economic policies swept the world. National barriers against the free mobility of capital were removed, privatization was seen as the magical solution to all economic problems from unemployment to inefficiency. To facilitate cooperation between universities and the industry, with the hope that such cooperation would boost the United State’s competitiveness in the knowledge economy especially against the rising “Asian Tigers” such as China and South Korea, the U.S. government passed a number of laws (see [Krimsky 2004](#), pp 30–31). The most important of these was the Bayh-Dole Act of 1980. This act gave small firms and universities the right to patent the results of publicly funded research. In 1987 the act was extended to cover big firms as well. The rationale behind these legal arrangements were purely commercial; they encouraged collaboration between universities and industry, more specifically, a technology transfer from the former to the latter.

Ideological: An ideology of neoliberalism accompanied these and similar arrangements. It was argued that a free, unregulated market economy was the most efficient mechanism for the allocation of resources. Accordingly, universities began to be seen as firm-like entities that needed to be guided by economic values such as efficiency, productiveness and profit. Universities were pushed to become entrepreneurial, and when coupled with the fear that their budgets would be cut due to economic concerns, they received the ideology positively.

Legal: A crucial Supreme Court decision in 1980 opened the gate for patenting both genetically modified living creatures and the genetic material itself. In the famous *Diamond v. Chakrabarty* case, the Supreme Court ruled with a 5–4 vote that artificially created organisms can be patented under the U.S. Patent Act. Thus, a patent was granted for a genetically engineered bacterium capable of breaking down crude oil. The majority opinion held that the bacteria was a useful “manufacture” not found anywhere in nature. The rest, as they say, is history. Soon after the Supreme Court decision, patents for DNA, RNA, proteins, cell lines, genes, genetic tests, gene therapy techniques, recombinant RNA techniques, genetically modified plants and even living animals were allowed by the U.S. Patent and Trademark Office (USPTO). By the year 2000, the USPTO issued about six thousand patents on genes, and about one sixth of them were human genes ([Krimsky 2004](#), p 66; for more about the role of the courts, see [Irzik 2007](#)).

Scientific: In the last several decades, we have also witnessed the revolutionary emergence of what might be called “technosciences”: computer science and technology, communication and information technologies, genetic engineering and biomedicine. Two features of technosciences strike the eye immediately: first, they blend science and technology in such a way that it is virtually impossible to make

a distinction between “pure” or “basic” science and “applied” science in these domains (hence the name “technoscience”); and, second, they hold the potential to respond to the demands of a globalized market by producing innovations that can bring generous profits. The technosciences became rapidly commercialized under the political, economic, legal and ideological conditions summarized above.

11.3 Benefits and Costs of Commercialization

On the surface everybody seems to benefit from the commercialization of academic science. Let us begin with the impact of the Bayh-Dole act. Prior to it, the U.S. federal government held approximately 30,000 patents, but only a very small part of it (roughly, 5%) led to any new products. The federal government simply did not have enough resources to convert the inventions into any commercial use. Through the act, it was hoped that universities, in collaboration with industry, would do what the federal government could not. Indeed, universities responded well; within less than two decades after the law was enacted, university-held patents increased tenfold, as contrasted with only a twofold increase in overall number of patents during the same period (Jasanoff 2005, p 235; see also Krinsky 2004, p 32 for further statistics on this issue). This brought financial (admittedly, modest) gains to the universities through royalties out of patents they hold or share. For example, in the year 2000 universities’ earnings from patent licensing totaled more than one billion dollars (Bok 2003, p 101).

Individual scientists, too, benefited from this situation as they enjoyed new opportunities to fund their researches and make money at the same time. While still holding their university positions and often being encouraged by the university administrators, many scientists became consultants, CEO’s or partners in these firms, others have started up their own companies, making literally millions of dollars (Kenny 1986; Krinsky 2004; Slaughter and Leslie 1997). Business firms were happy because they capitalized on the new inventions and increased their profits. Moreover, in return for the funds they offered to the universities, they enjoyed not only expert labor power, lab and equipment, but also prior or privileged access to the results of scientific research and shared or sole ownership of patents. Finally, it could be argued that the public also won because they benefited from new drugs and therapies that would otherwise not have occurred. In short, a miracle seems to have occurred.

There is, however, increasing evidence that this miracle has occurred at a considerable cost. The negative impact of commercialization on academic science can be seen both at the institutional and cognitive-epistemic level. Let us begin with the latter. Consider first the research problems and agendas. Generally speaking, these are shaped and given priority through a very intricate system that bears the marks of intrinsic theoretical interest and intellectual challenge, past scientific achievements, and the public benefit. The policies and developments outlined in the previous section resulted in a university-industry collaboration that skewed research toward what

is patentable and commercially profitable, especially in biomedicine, genetics and pharmacology. Research interests are increasingly shaped by commercial and corporate interests rather than by scientific value or social utility (Brown 2008). For example, there is little new research towards curing tropical diseases although millions of people, almost all of whom live in developing countries, suffer from them. “According to the World Health Organization, 95% of health related R&D was devoted to issues of concern primarily to the industrial countries, and only 5% to the health concerns of the far more populous developing world.” (World Development Report 1998, p 132) The reason for apathy seems to be that such research is just not sufficiently profitable.

In addition to the choice of research problems, commercialization seems also to affect the very content of scientific research in medicine. Several studies found a significant association between the source of funding and the outcome of scientific research. More precisely, they indicated that “private funding can bias the outcome of studies toward the interests of the sponsor” (Krimsky 2004, p 146). For example, an article published in the *Journal of General Internal Medicine* examined 107 controlled clinical trials which were classified along two dimensions: one, according to whether they favored a new or a traditional therapy, and, two, according to whether they were supported by a pharmaceutical manufacturer or by a nonprofit institution. The study found that 71% of the trials favored new therapies, and 43% of these were supported by a pharmaceutical firm. By contrast, of the 29% of the trials that favored the traditional therapies, only 13% were supported by pharmaceutical companies. Thus, there was a statistically significant association between the source of the support and the outcome of the research (Davidson 1986). Perhaps more tellingly, in none of the 107 cases examined a drug manufactured by the sponsoring company was found to be less effective than a rival drug manufactured by another firm! (For this and other examples, see Krimsky 2004, ch. 9 and Brown 2008).

One cause of such and other biases seems to be conflicts of interest, which can be defined as follows: “A researcher has a conflict of interest if and only if he or she has personal, financial, professional, or political interests that have significant chance of compromising the judgment of the average scientists in the conduct of research.” (Resnik 2007, p 111) Given that corporate-sponsored researches in the universities are increasing since the 1980s, we should expect to find an increase in the number of conflicts due to the *financial interests* of the scientists. Indeed, social researchers point out that financial conflicts of interests among scientists are relatively recent, and they have found a number of such cases (Krimsky 2004, ch. 8, Resnik 2007, pp 23–28). A typical case goes like this, as suggested by the preceding paragraph. Scientist A conducts a research that “shows” that drug D manufactured by company C is more effective than a rival drug R produced by another company. Later, it turns out that A has been sponsored by the manufacturer of D. An independent research by another scientist refutes the finding of the first research. Or, similarly, scientist A conducts a research that “shows” that drug D is effective in treating condition C. Later, it turns out that A has been sponsored by the manufacturer of D. Again, an independent research by another scientist refutes the finding of the first research.

Such cases are certainly interesting from a methodological viewpoint. Even if it is true that the biased outcome is indeed caused by the financial interest in question, as we all know, the existence of a correlation by itself is no “proof” of this. Are there then alternative explanations? For instance, could it be that journals are not much interested in publishing negative results, or might it be the case that drug companies support only those studies for which they have preliminary data favorable to them, as Koertge asks? (Koertge 2008; see also [Krimsky 2004](#), p 147) Which explanation is the best and simplest? These are exactly the kind of questions philosophers of science are well equipped to answer.

The examples of bias discussed above should be of interest to philosophers of science for another reason. Whether interests affect the content of science is a hotly debated issue in the philosophy of science. As is well known, the claim that it does and therefore the claim that the very content of science can be explained sociologically constitutes the cornerstone of the Strong Program in the Sociology of Scientific Knowledge. If it is true that financial interests are causing bias in scientific research in medicine, that would provide strong support for the Strong Program. Perhaps, then, its defenders should pay more attention to case studies in medicine than elsewhere.

Let me now turn to the impact of commercialization on the institutional aspects of science. Commercialization is threatening the social norms of science, what the famous sociologist Robert Merton has dubbed “the ethos of modern science”. By the term “scientific ethos”, Merton means the institutional values and norms that bind the community of scientists in their scientific research and activity. Merton lists four such norms: universalism, communalism, disinterestedness, and organized skepticism ([Merton 1973](#), pp 268–270). For lack of space, I will discuss only the first of these.

Communalism refers to the common ownership of scientific discovery or knowledge. Merton expresses it as follows: “The substantive findings of science are a product of social collaboration and are assigned to the community . . . Property rights in science are whittled down to a bare minimum by the rationale of scientific ethic. The scientist’s claim to ‘his’ intellectual ‘property’ is limited to that of recognition and esteem which, if the institution functions with a modicum of efficiency, is roughly commensurate with the significance of the increments brought to the common fund of knowledge.” (ibid., p 273) The rationale Merton has in mind is that new scientific knowledge always builds upon old knowledge and that scientific discoveries owe much to open and free discussion and exchange of ideas, information, techniques and even material (such as proteins). To be sure, there is competition, but it is mostly friendly and excludes collaboration seldom, if at all.

As we have seen in Section 11.2, as a result of the Supreme Court decision in *Diamond v. Chakrabarty*, genes, DNA, cell lines, and even living organisms like mice, whose genetic structure is sufficiently modified, became objects of intellectual property. It is no longer the case that “property rights in science are whittled down to a bare minimum by the rationale of scientific ethic”. This may very well be the reason why secrecy, which is the opposite of communalism, is spreading in an alarming way. When universities receive industrial support for their researches, they

sign protocols that often contain non-disclosure clauses that ban university scientists from publishing their findings without the written consent of the supporting company. In 1995 a study conducted by *New England Journal of Medicine* revealed that among the scientists in the top 50 universities receiving money from the US National Institute of Health, one out of four was involved in industry relationships and that they were twice as likely to engage in trade secrecy or to withhold information from their colleagues in comparison to those who were not involved in relationships with industry (Greenberg 2001, p 357). A recent study by Harvard Medical School reached similar conclusions. Forty-seven percent of geneticists reported that they were denied information, data, or materials related to published research results at least once in 3 years; 28% of them said that because of this they could not confirm the accuracy of published results (Krimsky 2004, p 83).

The relationship between the commercialization of academic science and the social norms of science have attracted the attention of many scholars (Brown 2008; Krimsky 2004; Resnik 2007; and especially Radder (ed.) forthcoming). Indeed, it would not be an exaggeration to say that Mertonian norms are going through a Renaissance after they were dismissed by some practitioners of social studies of science especially in the seventies (see, for example, Mulkay 1976 and Mitroff 1974). These critics argued that in practice scientists seldom acted in accordance with Mertonian norms, which in reality functioned as an ideology serving the interests of scientists, and that they even respected counter-norms. As Sergio Sismondi put it, “if there are both norms and counter-norms, then the analytical framework of norms does no work” (Sismondo 2004, p 26).

In a penetrating article, Hans Radder has responded to these criticisms. In particular, he has pointed out that the really interesting question is an ethical-normative one which goes beyond the narrow, descriptive concerns of the Strong Program: it is the question of “whether Merton’s ethos of science is a valuable perspective in an age of pervasive commodification of academic research” (Radder forthcoming, p 7). The answer, I think, is obvious.

Indeed, there has been a growing interest in the ethics of science, not just in Mertonian norms. Although the ethics of science owes much to Merton’s pioneering ideas, it contains other values such as scientific integrity, trust and openness not discussed by Merton. Moreover, the ethics of science goes beyond a general characterization of the ethos of science to cover specific codes of behavior. The topic of the ethics of science is receiving a great deal of attention not only by philosophers but also by scientists and scientific institutions. Many universities and institutions like National Institute of Health and scientific academies have established ethical codes of conduct for research. In 2007 the first World Congress on scientific integrity was held in Lisbon, organized by the European Science Foundation (ESF) and the U.S. Office of Research Integrity and presented its report (see the web page of ESF at <http://esf.org/index.php?id=4479>). In the report commercialization is especially mentioned as encroaching on academic science. Similarly, a 2003 Royal Society report with the sobering title “Keeping Science Open: The Effects of Intellectual Property Policy on the Conduct of Science” warned that there is evidence that “patenting can encourage a climate of secrecy that does limit

the free flow of ideas and information that are vital for successful science (see <http://royalsociety.org/document.asp?tip=0&id=1374>).

Thus, as Radder rightly points out, the present situation of commercialized academic research is very different from the one in which Merton wrote about the social norms of science in 1942 and also from the one in which the advocates of the Strong Program criticized him in the seventies (Radder forthcoming, p 9). Today, ethical codes of conduct have been pretty much institutionalized, and this very fact can be seen as a reflection of a sensitivity to the unease caused by commercialization. Indeed, the framework of values and norms such as communalism, scientific integrity and openness can and does function as at least a partial shield against its negative effects, doing its “work”.

11.4 Discovery Versus Invention, Fact Versus Artifact

Customarily, we think of the distinctions between discovery and invention and between fact and artifact as follows: while facts are discovered, artifacts are invented; whereas facts belong to the domain of nature, artifacts belong to the domain of culture. Thus, the concept of discovery applies to entities like planets and electrons, phenomena such as blackbody radiation and the Compton effect, facts like $E = mc^2$; by contrast, things like microscopes, air pumps, radios and atomic bombs are invented, they are artifacts. Through discovery secrets of nature are disclosed, through invention new objects that did not exist in nature before are created by human ingenuity and skill. All of this suggests that the distinctions in question are ontologically grounded.

However, the situation is a lot more complicated than this. To see this, look at some “hard” cases (Resnik 2002). Genes are clearly part of our biological make-up, but they do not occur in nature in a pure and isolated form; they must be removed from their chromosomes, an activity which naturally requires much ingenuity and skill. Are genes then discovered or invented items? Or consider the Harvard “oncomouse”, a genetically modified animal that is made susceptible to cancer. While mice are certainly natural creatures some of which may develop a predisposition to cancer due to natural mutations, the Harvard oncomouse is “created” by Harvard scientists by genetic engineering. Again, is the Harvard oncomouse an invention or not? All of it or only a part of it? As David Resnik points out, “I think people most people would agree that a person who carves out a flute out of a stick of wood invents part of the item but not the whole item. One part of it – its design – is a human invention, but another part – its material – is not. If the whole flute is not an invention, then does it make sense to say that the whole mouse is an invention?.” (Resnik 2002, p 144).

David Resnik has also argued that in such borderline cases whether something is an invention or not is a pragmatic matter that depends more on human purposes and values than on ontology or metaphysics. These values may be scientific, technological, religious, moral, economic or legal, which may sometimes conflict with one

another. Here is then an interesting set of questions for philosophers of science: is there a way of reducing the hard cases to matters of ontology? If not, which values should decide the issue and when they conflict, how should they be weighed? How do scientists themselves conceptualize their findings in borderline cases: as discoveries or as inventions? Is there a change in their outlook in the last several decades?

How we draw the discovery-invention distinction has a direct bearing on patenting practices: whereas inventions can be patented according to the patent laws in the U.S. and in many other countries, discoveries cannot. If the discovery-invention distinction is a matter of ontology, then it follows that whether something can be patented or not is also a matter of ontology and therefore objectively decidable. But if the distinction is a pragmatic (“socially constructed”, as social constructivists might say) one at least so far as some cases are concerned, then the discussion shifts to the domain of values and purposes.

As it turns out, the U.S. patent office did grant a patent for the whole of Harvard oncomouse in 1988, and as we saw earlier genes and many similar items are also being routinely patented since the 1980 Supreme Court decision. This has another striking consequence that should be of interest to philosophers of science. Whatever commercial benefits it may provide, patenting of such life forms diminishes the space of intellectual commons. As Sheldon Krimsky put it, “the upshot of this decision [of the United States Patent and Trademark Office to patent genes] has made every gene sequencer an ‘inventor’ or ‘discoverer of patentable knowledge,’ which has inadvertently thrust normal genetic science into entrepreneurship and basic biological knowledge into a realm of intellectual property.” (Krimsky 2004, pp 69–70) Thus, what used to belong to the realm of public knowledge becomes private property for a period of time (often 20 years), excluding others from using it or requiring them to pay for it when they want to use it. No doubt, patents can stimulate scientific/technological innovations, but they can also hinder the development and progress of science since new knowledge always builds upon the old one.

11.5 Concluding Remarks

I have argued that commercialization of academic science and commodification of scientific knowledge more specifically has a number of effects on science, some good and others plainly undesirable. These effects range from the choice of scientific problems to the content of science, from the discovery-invention distinction to the ethos of science, all of which should be of interest to philosophers of science in one capacity or another. I would like to conclude by drawing attention to a final global worry I have.

Commercialization of academic science has the potential of subverting science’s cognitive and social functions. Science is held in high esteem by the public precisely because it has delivered what it is expected of it. People generally have confidence in the findings of science, trust the scientists’ judgments especially in matters of health and environment, and count on their independent critical voice. The image

of a scientist who is secretive, partial, and interested more in money than in truth or social utility is destructive of the social status of science. Such an image may erode public confidence in the results of science and undermine science's social legitimacy. Anyone who cares for science cannot and should not remain indifferent to such a disastrous possibility.

Acknowledgements I acknowledge the support of the Turkish Academy of Sciences.

References

- Bok D (2003) *Universities in the market place*. Princeton University Press, Princeton, NJ
- Boyle J (1997) *Shamans, software, and spleens*. Harvard University Press, Cambridge
- Brown JR (2008) Community of science (R). In: Carrier M, Howard D, Kourany J (eds) *The challenge of the social and the pressure of practice: science and values revisited*. University of Pittsburgh Press, Pittsburgh, pp 189–216
- Davidson RA (1986) Source of funding and outcome of clinical trials. *J Gen Int Med* 1:155–158
- Etzkowitz H, Webster A (1995) Science as intellectual property. In: Jasanoff S, Markle GE, Petersen JC, Pinch T (eds) *Handbook of science and technology studies*. Sage, Thousand Oaks, CA, pp 480–505
- Greenberg DS (2001) *Science, money, and politics*. The University of Chicago Press, Chicago, IL
- Irizik G (2007) Commercialization of science in a neoliberal world. In: Buğra A, Ağartan K (eds) *Reading Polanyi for the twenty-first century*. Palgrave MacMillan, New York, pp 135–154
- Jasanoff S (2005) *Designs on nature*. Princeton University Press, Princeton, NJ
- Kenny M (1986) *Biotechnology: the university-industrial complex*. Yale University Press, New Haven, CT
- Koertge N (2008) Expanding philosophy of science into the moral domain: response to Brown and Kourany. *Philos Sci* 75:779–785
- Krimsky S (2004) *Science in the private interest*. Rowman & Littlefield, Lanham, MD
- Magnus D, Kaplan A, McGee G (eds) (2002) *Who owns life?* Prometheus Books, Amherst
- McSherry C (2001) *Who owns academic work?* Harvard University Press, Cambridge
- Merton R (1973) *The sociology of science*. The University of Chicago Press, Chicago, IL
- Mirowski P, Sent E-M (2008) The commercialization of science and the response of STS. In: Edward JH, Amsterdamska O, Lynch M, Wajeman J (eds) *The handbook of science and technology studies*, 3rd edition. MIT, Cambridge, pp 635–689
- Mitroff I (1974) Norms and counter-norms in a select group of the Apollo moon scientists: a case study of the ambivalence of scientists. *Am Sociol Rev* 39:579–595
- Mulkay M (1976) Norms and ideology in science. *Social Sci Inform* 15:637–656
- Radder H (ed) (forthcoming) *The commodification of academic research: analyses, assessments, alternatives*. University of Pittsburgh Press, Pittsburgh
- Radder H (forthcoming) Mertonian values, scientific norms, and the commodification of academic research. In: Radder H (ed) *The commodification of academic research: analyses, assessments, alternatives*. University of Pittsburgh Press, Pittsburgh
- Resnik D (2002) Discoveries, inventions, and gene patents. In: Magnus D, Caplan A, McGee G (eds) *Who owns life?* Prometheus Books, Amherst
- Resnik D (2007) *The price of truth*. Oxford University Press, Oxford
- Sismondo S (2004) *An introduction to science and technology studies*. Blackwell, Malden
- Slaughter S, Leslie L (1997) *Academic capitalism*. The John Hopkins University Press, Baltimore, MD
- World development report (1998) World Bank. Oxford University Press, New York

Chapter 12

Some Consequences of the Pragmatist Approach to Representation

Decoupling the Model-Target Dyad and Indirect Reasoning

Tarja Knuuttila

12.1 Introduction

In an interesting recent effort to specify the distinct nature of modeling Michael Weisberg (2007) and Peter Godfrey-Smith (2006) argued that what distinguishes modeling from other types of theory construction is the strategy of indirect representation and analysis it makes use of. By this they mean that instead of directly striving to represent some aspects of real target systems, modelers seek to understand the real world through the procedure of constructing and analyzing hypothetical systems, in other words models. Thus they posit that modeling constitutes a specific *theoretical practice*, something that has escaped the notice of many philosophical accounts concerning the interrelationships between theories and models.

Whereas Weisberg focuses on explicating in detail what modeling as indirect representation and analysis consists of, Godfrey-Smith approaches it also from a wider perspective, “as an approach with both strengths and weaknesses, with effects on the sociology of science and perhaps with a distinctive historical signature” (2006, 726). I find both proposals feasible and intuitively very much to the point as regards modeling practice. However, even though the notion of indirect representation constitutes the core of modeling as a distinct theoretical practice, neither Weisberg nor Godfrey-Smith really attempts to relate his views on indirect representation to the recent discussion on scientific representation.¹ This is understandable given that they are both first and foremost interested in the nature of modeling per se. Yet, as the burgeoning discussion on scientific representation has taken place exactly in the context of models, I think it would be worthwhile to study how the notion of indirect representation relates to it.

T. Knuuttila (✉)

P.O. Box 24 (Unioninkatu 40 A), SF – 00014 University of Helsinki, Finland
e-mail: ttknuut@mappi.helsinki.fi

¹ Although Godfrey-Smith (2006) discusses Giere’s (1988, 1999) views on models and representation, and Weisberg in turn goes quickly through several recent positions taken in the debate on scientific representation, settling for the observation that no consensus has emerged in that discussion.

In the following I will show how Weisberg and Godfrey-Smith effectively decouple through their thesis of indirect representation the model-target dyad, which has been the constitutive unit of analysis² in the discussion on models and scientific representation. What is more, I will show how a similar conclusion has been drawn in the discussion on scientific representation, although on different grounds. Taken together, these positions challenge us to think anew the ways in which models give us knowledge and enable us to learn from them about real-world systems. To this effect, I will suggest that modeling as *results-driven activity* gives us knowledge through *indirect reasoning*, with which I refer to the way modelers proceed via the output and results of their models to consider the underlying mechanisms that might have produced the phenomena of interest.

12.2 The Thesis of Indirect Representation

In his article “Who is a Modeler” Weisberg (2007) redirects the focus from models to the activity of modeling, suggesting that modeling proceeds in three stages. Firstly, a model is constructed, after which and secondly, the modeler proceeds to refining, analyzing and articulating its properties and dynamics. It is not until the third stage that the relationship between the model and any target system is assessed, “if such an assessment is necessary” (2007, 209). Godfrey-Smith, in turn, offers two stages or “moves”. The first is that of “specification and the investigation of the hypothetical system”, i.e. the model. Like Weisberg he claims that in model-based science the “resemblance relations” between the model and the real systems are typically first considered in the second stage – although this stage is often left implicit as the modelers may go on studying the model systems created without too much explicit attention to their relationship with the world.

The claim that model construction happens *before* the possible real target systems are considered runs counter to the conventional philosophical understanding of models as representations of some target systems, an idea that has motivated recent discussion on scientific representation, as I will argue further below. More often than not, target systems are understood in terms of real world systems. This being the case, the burden of proof lies on the shoulders of Weisberg and Godfrey-Smith. If models are not representations of some real target systems at the outset, *what* is represented in them and *how* that is supposed to happen? What, then, is indirect representation all about? Interestingly, neither Weisberg, nor Godfrey-Smith tries to define indirect representation. In trying to specify the characteristics of indirect representation both authors rather revert to scientific examples and their observations concerning them.

² I am indebted to Paul Humphreys (2004) the idea of the unit of analysis: in an insightful way he has applied this notion, which plays an important role in the methodology of the social sciences, to the analysis of the computational science.

Weisberg contrasts Vito Volterra's style of theorizing, which he takes as an example of modeling, to "abstract direct representation" as exhibited by Dimitri Mendeleev's Periodic Table. According to Weisberg, Volterra studied the special characteristics of post-World-War-I fish populations in the Adriatic Sea by "imagining a simple biological system composed of one population of predators and one population of prey" to which he attributed only a few properties, writing down a couple of differential equations to describe their mutual dynamics. The word "imagining" used by Weisberg here is important since it captures the difference between the procedures of direct and indirect representation. He stresses the fact that Volterra did not arrive at these model populations by abstracting away properties of real fish, but rather constructed them by stipulating certain of their properties (210). Unlike Volterra, he claims, Mendeleev did not build his Periodic Table via the procedure of constructing a model and then analyzing it. In developing his classification system he was rather working with abstractions from data in an attempt to identify the key factors accounting for chemical behavior. Thus, in contrast to modelers such as Volterra, he was trying to "represent trends in real chemical reactivity, and not trends in a model system" (215, footnote 4).

Godfrey-Smith (2006) studies more recent examples, focusing on two influential books on evolutionary theory: Leo Buss's *The Evolution of Individuality* (1987) and Smith and Szathmáry's *The Major Transitions in Evolution* (1995). For him they represent an ideal example being written about at the same time and on partly overlapping topics. In his study on the evolution of multicellular individuals from the lower-level competition on the level of cell lineage, Buss examined the "actual relations between cellular reproduction and whole-organism reproduction in known organisms" (2006, 731). As opposed to Buss's approach, Smith and Szathmáry describe "idealized, schematic causal mechanisms". Rather than studying actual systems they engage in modeling, that is in examining "tightly constrained" possible – or fictional – systems. Thus their explanations "would work just as well in a range of nearby possible worlds that happen to be inhabited by different organisms", which endows them with what Godfrey-Smith aptly calls "modal 'reach'" (2006, 732).

The crucial difference between abstract direct representation and indirect representation does not concern whether one abstracts or approximates, selects or even idealizes. Scientific representation involves all these, but in engaging in such activities modelers do not even pretend to be primarily in the business of representing any specific real system. For them the models come first. The distinguishing feature of the "strategy of model-based science" is that the modelers do not try to identify and describe the actual systems, but proceed by describing another more simple hypothetical system (Godfrey-Smith 2006). Thus model-based science could be characterized by the "deliberate detour through merely hypothetical systems" of which it makes use (2006, 734).

Consequently, it follows from the thesis of indirect representation that models should be considered independent objects in the sense of being independent from

any real target system.³ As such, I argue, it means a departure from the representational paradigm, which has taken the *model-target dyad* as a basic unit of analysis concerning models and their epistemic value. Even though this is not the specific goal Weisberg and Godfrey-Smith set themselves, it is a clear consequence of their approach. Interestingly, but for quite different reasons, the recent discussion on models and representation has also led to the same conclusion. A look at this discussion will shed light on the reasons why.

12.3 Models and Scientific Representation

The rather striking feature of the discussion on scientific representation is that it has been, so far, conducted almost exclusively in the context of modeling. This may seem curious given that scientific endeavor employs manifold representations that are not readily called models. Such representations include visual and graphic displays on paper and on screen, such as pictures, photographs, audiographic and 3D images, as well as chart recordings, numerical representations, tables, textual accounts, and symbolic renderings of diverse entities such as chemical formulas. One rationale for this discussion, apart from the specific historical reasons for it – a topic I will not touch upon here – is that models have traditionally been taken to be representations. This conviction is of far more distant origin than the semantic approach to models in its various guises, which until the recent decade has been the dominant approach to models.

What has been characteristic of the recent discussion on models and representation is the double move of treating models as representations *and* ascribing their epistemic value to representation. Accordingly, what several philosophers have taken as their task is to give an account of *scientific* representation, given that it is generally agreed that models give us knowledge because they are able to represent some real target systems (e.g., French 2003; Giere 2004; Suárez 2004; Contessa 2007; Mäki forthcoming; Frigg forthcoming). The basic unit of these accounts has been the model-target dyad, and the question has concerned the kind of relationship between a model and its real target system by virtue of which the model can give us scientific knowledge.

The most straightforward answer to the question of representation has been given by the semantic or structuralist accounts. These accounts have focused on the properties or features that models supposedly share with their target systems, thus concentrating solely on the model-target dyad. According to the semantic or structuralist notion of representation, models specify structures that are posited

³ Other authors have also recently suggested that models could be conceived of as independent objects, although by this they mean different things. Morrison and Morgan (1999) conceive of models as partly independent of theory and data. Knuuttila (2005) treats them as independent entities in the sense of loosening them from any predetermined representational relationships to real target systems. This comes close to what Weisberg and Godfrey-Smith mean by independence.

as possible representations of either the observable phenomena or, even more ambitiously, the underlying structures of the real target systems. This relation of representation between a model and its target system has been formulated in terms of isomorphism, partial isomorphism, or – less frequently – similarity (e.g., van Fraassen 1980; French and Ladyman 1999; da Costa and French 2003; Giere 1988). Thus, according to the semantic view, the structure specified by a model represents its target system if it is either structurally isomorphic or somehow similar to it.

Pragmatist critics of the semantic conception have argued, among other things, that isomorphism – being a symmetric, reflexive and transitive relation – does not satisfy the formal and other criteria we might want to affirm of representation (see e.g., Suárez 2003 and Frigg 2003, of whom Suárez has extended this critique also to similarity). For instance, both isomorphism and similarity denote a symmetric relation, whereas representation does not: we want a model to represent its target system but not vice versa. Moreover, the isomorphism account does not accept false representations as representations. The idea that representation is either an accurate depiction of its object – which is interpreted in terms of isomorphism within the structuralist conception – or it is not representation at all does not fit our actual representational practices. These problems appear to be solved once the pragmatic aspects of representation are taken into account. The users' intentions create the directionality needed to establish a representative relationship: something is being used and/or interpreted as a model of something else, which makes the representative relation triadic, involving human agency. This also introduces indeterminateness into the representative relationship: human beings as representers are fallible.

In stressing the importance of human agency for what representation is all about, the pragmatic approaches criticize the assumption of the semantic conception that representation is a dyadic relation of correspondence between the representational vehicle (a model) and its target (Suárez 2004; Giere 2004). The dyadic conceptions attempt, as Suárez has put it, “to reduce the essentially intentional judgments of representation-users to facts about the source and target objects or systems and their properties” (2004, 768). Thus Suárez spells out that what is actually at stake is whether or not the possibility of representation can be based on some privileged parts or properties that the actual representative vehicles are supposed to share with their target objects.

Even though the basic problem of representation that the pragmatist approaches have set out to solve has been cast out in terms of the model-target dyad, their analyses in fact decouple that dyad by introducing representation users and their intentions and purposes. Thus, the outcome of this discussion fits in well with the idea of indirect representation, according to which models, being independent objects, “do not have a single, automatically determinable relationship to the world” (Weisberg 2007, 218). It is also worth noting that the question of fiction and the ontology of models has started to interest pragmatists of representation in particular (e.g., Frigg forthcoming; Suárez 2008), which resonates well with both Weisberg's and Godfrey-Smith's views.

However, the weight given by both Weisberg and Godfrey-Smith to similarity (or resemblance) concerning how models give us knowledge and understanding seems questionable in the light of the pragmatist view on representation. The problem is not only about the vagueness of the notion of similarity – which point is habitually noted in this context – but also that it does not accomplish much from the philosophical point of view. Namely, by invoking the notion of similarity Weisberg and Godfrey-Smith are implicitly taking a stand on the issue of representation, and whereas the way they loosen the model-target dyad is something that pragmatists of representation would agree on, here the roads divide. For instance Weisberg puts it bluntly that in order for us to learn about the real world “the model must be *similar* to a real-world phenomenon in certain appropriate respects” (2007, 218). Currently, it seems that those engaged in the discussion on scientific representation are not willing to endorse the similarity account without also reverting to users. A good example is provided by Ronald Giere, the most well known proponent of the similarity account, who instead of arguing for similarity prefers to account for representation in terms of an “intentional account of representation” (Giere forthcoming).

It seems to me that the reason why evoking mere similarity in an effort to establish a representational relationship between the model and a real-world target system is problematic, apart from the arguments already referred to, lies in its observer-dependent nature. If it is a case that many if not most things can be taken to be similar to most other things, then it is we who pick the “appropriate similarities” – and in this sense Giere’s turn from the similarity account of representation to an intentional account seems an entirely appropriate step to take. Indeed, this observation was already present in Giere’s classic 1988 account, in which he did not appear too worried about the vagueness of the notion of similarity, claiming that cognitive sciences are accumulating evidence that “human cognition and perception operate on the basis of some sort of similarity metric” (1988, 81). Thus similarity has its proper place in our cognitive endeavor, but not the place to which it is habitually relegated in the discussion on representation. The point is that it tells more about our cognitive functioning than specifically about the epistemic value of modeling. We may tend to recognize similarities between different things, but that does not yet make the similarities in question epistemically interesting. Consequently, even though similarity considerations undoubtedly play a part in our cognitive judgments, they do not take us very far in understanding how we learn from models.

Apart from decoupling the model-target dyad, the pragmatist accounts of representation have also some other consequences worth of mentioning. Namely, once we introduce users into the relationship of representation, its explanatory power starts to fall apart. The gesture of relating representation to the intentional activity of model users solves many problems of the semantic notion, but this comes at a price: if representation is grounded primarily in the specific goals and the representing activity of humans as opposed to the properties of the representative vehicle and its target, nothing very substantial can be said about it in general. This has been explicitly admitted by proponents of the pragmatic approach (cf. Giere 2004), of whom Suárez (2004) has gone farthest in arguing for a “deflationary”, or minimalist, account of representation that seeks not to rely on any specific features that

might relate the representational vehicle to its target. The minimalist approach has rather radical consequences in terms of how the epistemic value of models should be conceived of. Namely, if we attribute the epistemic value of a model to its being a representation of some target system and accept the minimalist pragmatic notion of representation, not much is established about how we can learn from models. This naturally raises the question of whether there is any other way to approach models that could give us some more insight into their epistemic functioning. Before I go into this question allow me to make still one more point concerning the relationship of the thesis of indirect representation to the general discussion on scientific representation.

As far as the relationship of representation is concerned, the thesis of indirect representation divides it into two parts: to the *construction* of models and to the *use* of them. I find this distinction an important contribution to the discussion on scientific representation. Since the model-target dyad has been taken as the starting point of the analysis, no such distinction has been made so far in this discussion. One reason for this is the frequently tacit assumption that models are inherently *models of* some pre-established target systems. They are taken to depict some real-world target systems *at the outset* and thus the question of representation concerns the conditions under which a model succeeds in representing the target (given that it is also assumed that representation is a condition for our learning from models). One common idea behind this line of reasoning is that models typically isolate some causal factors or tendencies of a system of interest and abstract away from other disturbing factors by means of suitable idealizations (e.g., Cartwright 1998; Mäki 2005). However, according to the thesis of indirect representation, the model need not be bound in this way to a real-world system. Even though the model construction makes use of available theoretical and empirical knowledge, this knowledge is mediated by the construction of a simpler *imagined* system. In the following I will suggest that the same characteristic detour also applies to our learning from models. In this case too, the links forged with real-world systems are looser and more complicated than the mere appeal to similarity suggests. Thus, from the perspective of scientific practice, the knowledge and understanding gained via modeling are achieved through various kinds of inferences derived from models combined with various kinds of background knowledge and other evidence.

12.4 Results-Drivenness in Modeling and Indirect Reasoning

Given both the thesis of indirect representation *and* the minimalist pragmatist account of representation the crucial question is how models as independent hypothetical objects enable us to understand and learn about the world. In order to answer this question let me to consider once again the assumed similarity of models to real-world target systems, this time not as a general answer to the question of representation but rather from the perspective of the practice of modeling. In this respect, two relevant observations arise as regards the further features of modeling

as a specific theoretical activity. Firstly, in considering models as imaginary entities both Weisberg and Godfrey-Smith note how they frequently concern non-existing systems such as three-sex biology. In these cases the modelers are clearly trading with fiction, in other words dealing with the possible and the non-actual. Indeed, scientific models typically provide exemplifications of the functioning of ideal, schematized mechanisms, as well as how-possibly and what-if-things-were-different types of explanations. It seems to me far from clear what kind of similarity comparisons between these modeled imaginary and non-existent systems and the real-world ones we are supposed – or even able – to make. This, in turn, is bound to lead one to ask, secondly, what sort of similarity appraisals are inherent in modeling.

I suggest that modeling is fundamentally a *results-driven* theoretical activity in which surrogate hypothetical systems, or models, are constructed keeping in mind the effects they are supposed to produce⁴. As models are typically valued for their performance and their results or output, the relevant similarities that modelers are primarily after are those between the model output and some stylized features of the phenomena of interest. The way interesting models are also expected to produce a priori unexpected results or to account for different empirical findings that, according to earlier theoretical knowledge, have been considered to be contradictory also point to the results-driven nature of modeling. Furthermore, it is backed up by the systemic holistic character of models, which distinguishes them from many other scientific representations that often fragment and analyze an object or specimen to its further details.

From this perspective, I suggest, that modelers engage in *indirect reasoning* by making use in their knowledge acquisition the results derived from purposefully designed hypothetical systems. Thus indirect reasoning makes a natural companion for indirect representation. Instead of directly trying to represent some selected aspects of a given real target system – as has conventionally been assumed – modelers proceed in a roundabout way, seeking to build hypothetical systems in the light of their *anticipated results* or of certain general features of phenomena they are supposed to exhibit. If a model succeeds in producing the expected results, i.e. some features of the phenomena of interest, it provides an interesting starting point for further theoretical conjectures and inferences, concerning the underlying real mechanisms, for instance.

This results-orientation also accounts for why modelers frequently use the same cross-disciplinary computational templates (Humphreys 2004), such as well-known general equation types, statistical distributions and computational methods. The overall usability of computational templates is based on their generality and the observed similarities between different phenomena. Thus there is an element of opportunism in modeling: the template that has proven successful in producing certain features of some phenomenon will be applied to other phenomena, often studied within a totally different discipline. This is certainly true of Lotka-Volterra equations, the example cited by Weisberg, which have been used in disciplines as

⁴ In his work on robustness Weisberg has targeted the epistemic importance of model results and the ways of guaranteeing their generality (e.g., Weisberg 2006).

different as biology, ecology, chemistry, physics and economics. In these areas they are typically applied to phenomena that usually exhibit complex fluctuations. It is also telling that the Lotka-Volterra model had a renaissance in the 1970s in the context of chaos and complex systems theory, when researchers became interested in exploring the nonlinear dynamics of the model. Last but not least, the aim of getting the model to bring forth results also explains why tractability considerations frequently override the search for realistic representation.

Looking at models from the perspective of their results-orientedness, I suggest, explains the very interest modelers have in their properties and dynamics, but it also accounts for their important instrumental uses in prediction, for instance, which also relies on the results they produce. Moreover, this perspective avails itself of different cognitive strategies: it is not limited to simplified mathematical models, but also takes in simulations in which output representations are crucial in terms of creating pragmatic understanding oriented towards control, design rules and predictions. In fact, simulations have been considered problematic in terms of the representational understanding of models because “instead of creating a comprehensive, though highly idealised, model world, [they] squeeze out the consequences in an often unintelligible and opaque way” (Lenhard 2006, 612).

The epistemological justification of indirect representation and indirect reasoning lies, as I see it, in contesting the traditional representational view that assumes that we already knew the relevant systems or causal mechanisms to be represented, and had the suitable representational means at hand for doing it. As far as scientific practice is concerned, this is hardly the case, which also accounts for the characteristic modal nature of modeling: the very interest of modelers is also studying different non-actualized and inexistent systems in an effort to chart various possibilities and thus to gain further understanding of the phenomena in question. This, in turn, has consequences for how models as entities should be understood. I take it that the path to the imaginary goes through the concrete and the manipulable, in other words through the model description, which allows modelers to experiment with various possibilities. Thus the material, concrete dimension of models embodied in some representational medium is crucial to their epistemic functioning.

12.5 Conclusion

Both the thesis of indirect representation and the pragmatic notion of representation point out that the ties between models and real-world systems are looser than is customarily assumed. The specific contribution of the former lies in its distinguishing of two aspects in the relationship of representation, the construction of models on the one hand and their use on the other. This shifts the focus from the model-target dyad to the very *activity* of modeling. As far as the discussion on scientific representation is concerned, however, the appeal of both Weisberg and Godfrey-Smith to similarity (or resemblance) with regard to our learning from models seems somewhat too hasty and would need more fine-grained analysis. Towards this end I have suggested that

not just model construction but also inferring from models (i.e. their use) is indirect, proceeding through the results they give to consider the possible underlying mechanisms. Studying the model results under various assumptions – and in relation to other models and evidence, allows further inferences concerning the actual and the possible, which is why I have called this specific kind of model-enabled reasoning *indirect reasoning*.

References

- Cartwright N (1998) Capacities. In: Davis JB, Wade Hands D, Mäki U (eds) *The Handbook of economic methodology*. Edgar Elgar, Cheltenham, pp 45–48
- Contessa G (2007) Representation, interpretation, and surrogative reasoning. *Philos Sci* 71:48–68
- da Costa NCA, French S (2003) *Science and partial truth*. Oxford University Press, New York
- French S (2003) A model-theoretic account of representation (or, I don't know much about art . . . but I know it involves isomorphism). *Philos Sci* 70(Proceedings):1472–1483
- French S, Ladyman J (1999) Reinflating the semantic approach. *Int Stud Philos Sci* 13:103–121
- Frigg R (2003) *Representing scientific representation*. Ph.D. dissertation, London School of Economics, London
- Frigg R (forthcoming) Models and fiction. *Synthese* (forthcoming 2009) DOI: 10.1007/s11229-009-9505-0
- Giere RN (1988) *Explaining science: a cognitive approach*. The University of Chicago Press, Chicago, IL/London
- Giere RN (2004) How models are used to represent reality. *Philos Sci* 71(Symposia):742–752
- Giere RN (forthcoming) An agent-based conception of models and scientific representation. *Synthese* (forthcoming 2009) DOI: 10.1007/s11229-009-9506-z
- Godfrey-Smith P (2006) The strategy of model-based science. *Biol Philos* 21:725–740
- Humphreys P (2004) *Extending ourselves. Computational science, empiricism and scientific method*. Oxford University Press, Oxford
- Knuuttila T (2005) Models, representation, and mediation. *Philos Sci* 72(Proceedings):1260–1271
- Lenhard J (2006) Surprised by a nanowire: simulation, control, and understanding. *Philos Sci* 73(Symposia):605–616
- Mäki U (2005) Models are experiments, experiments are models. *J Econ Method* 12:303–315
- Mäki U (forthcoming) Models and the locus of their truth. *Synthese* (forthcoming 2009) DOI: 10.1007/s11229-009-9566-0
- Morrison M, Morgan MS (1999) Models as mediating instruments. In: Morgan MS, Morrison M (eds) *Models as mediators. Perspectives on natural and social science*. Cambridge University Press, Cambridge, pp 10–37
- Suárez M (2003) Scientific representation: against similarity and isomorphism. *Int Stud Philos Sci* 17:225–244
- Suárez M (2004) An inferential conception of scientific representation. *Philos Sci (Symposia)* 71:767–779
- Suárez M (2008) Scientific fictions as heuristic rules of inference. In: Suárez M (ed) *Fictions in science: philosophical essays on modeling and idealization*. Routledge, London, pp 158–178
- van Fraassen B (1980) *The scientific image*. Oxford University Press, Oxford
- Weisberg M (2006) Robustness analysis. *Philos Sci* 73(Symposia):730–742
- Weisberg M (2007) Who is a modeler. *Br J Philos Sci* 58:207–233