Participants Papers

REVISITING THE NEURO-TURN IN THE HUMANITIES AND NATURAL SCIENCES

DAVID BUDTZ PEDERSEN
University of Copenhagen

ABSTRACT: In the last twenty years there has been an unprecedented advance in neuroscience, in fields such as pharmacology, neurology and behavioural genetics. A growing number of ethicists, legal scholars and philosophers have begun to analyse the implications of these advances; from the use of fMRI scanning to the way in which scientific disciplines are organised in order to deliver techno-economic innovation. Drawing on recent contributions in philosophy of science and science policy studies I argue that much like the optimism surrounding the benefit of advances in pharmacogenomics and gene therapy, there is a risk of overstating the real impact and explanatory power of advances in the neurosciences.

1. Introduction

In the last twenty years there has been an unprecedented advance in the neurosciences, in fields such as pharmacology, neurology and behavioural genetics. A growing number of ethicists, social scientists, legal scholars and philosophers have begun to analyse the impacts of these advances, from the medical benefits and risks of the increasing prescription of psychotropic drugs to the ways in which scientific disciplines are organised to make them serve techno-economic innovations (Schneider, 2010). Drawing on recent contributions in philosophy of science and science policy studies, this paper suggests that much like the situation surrounding the advances in pharmacogenomics, there is a trend of overstating the real impact of advances in the neurosciences (Brown & Michael, 2003) [1].

While a number of scholarly projects have recently emerged to analyse the challenges and pitfalls of the «neuro-turn» in the humanities and sciences, the ambition of this paper is to take a step back and to probe deeper into the effects and causes of the ongoing neuroscientific revolution. Hence the objective is not one of debunking scientific research en bloc, but rather to raise the awareness of how mechanisms enforced from outside science possibly stimulate an explanatory race in which scientists are stipulating ever more radical claims and explanations. It seems that an increased pressure to apply for funding (in competition with other researchers, disciplines and universities) creates an unbalanced emphasis on promising scientific breakthroughs at the expense of less promissory disciplines, such as the social sciences and humanities. More fatally, the pressure to apply for funding and forecast entire research programmes in terms of expected future scientific findings turns research into an explanatory arms race; the fundamental structure of academic behaviour is changed, resulting in a self-defeating and hence counterproductive division of labor (Laudel, 2006; Landes, Andersen & Nielsen, 2011).

Without advocating the abandonment of the principle of competition in science, I find it regrettable that the competitive imperative in the current science system is fostering

[1] In the following, I conceive neuroscience broadly as to include a diverse set of disciplines ranging from neurobiology to cognitive psychology and anthropology. In between are neurophysiology, neuropsychology, neurolinguistics, artificial intelligence and cognitive sociology.

a number of unintended side-effects. In particular current trends in the way science is funded and evaluated are likely to change the relationship between the natural sciences and the humanities. Long-standing debates about the «the hard problem» of consciousness and associated ontological questions are today sidetracked by reductionist research programmes which has as their goal to eliminate more pluralist models of consciousness, intentionality and human behaviour (Burge, 2010).

The relation between brain and mind, to take one example, have been the critical topic of scientific and philosophical enquiry at least since the Renaissance. However, the intensity of this debate has risen sharply in the past decade as the scientific techniques for studying the functioning of the brain and the range of technologies for making more or less subtle alterations to its performance have expanded (Ashcroft & Martin, 2005). For example, neuroimaging techniques such as magnetic resonance imaging has gained significant visibility in public media and science policy reports over the past ten years, as demonstrated by Racine *et al*. (2006). In this sense brain imaging studies represent the trend and expectations surrounding current neuroscience, not because these studies are devoid of scientific substance but because the claims made for them exceed the explanations they actually provide. Neuroimaging and other neurotechnologies have become increasingly representative of «real» cognition as opposed to intentions, emotions, consciousness, belief and folk psychology (Weisberg *et al*., 2008; Gerrans, 2009).

Perceived in isolation the phenomenon of promoting certain scientific technologies or epistemic models as particularly intellectual promising might be somewhat harmless. It might even be plausible to comprehend such explanatory aspirations as an intrinsic aspect of all scientific endeavors. But seen through the lens of the current zero-sum-game of scientific funding and competition the growing epistemic priority of the neurosciences in understanding human behaviour is likely to affect the distribution of resources and prestige among disciplines. In this perspective, the attempt to promote a more moderate methodological stance based on a pluralist methodology, which has been at the heart of philosophical debates on model-driven science in the past years, is possibly undermined by practical incentives to overstate and oversell the scientific impact of neuroscience.

The argument of the present article is that the changing economic and political context of science is a critical issue in understanding the current «epistemic drift» towards neuroscience (Elzinga, 2004). Across U.S. and Europe, the promotion of neuroscience is part of a larger shift in universities, such as growing commercialization and new funding mechanisms. Whether or not the promises will be realised, research funding is embedded with expectations of new therapies, societal solutions and economic growth. Neurotechnology such as brain-based devices, drugs and diagnostics is expected to be a $145 billion industry already by 2025. Hence, as noted by Robinson (2010), «... neuroscience requires and invites a way of thinking about science in relation to its capacity for translating into other agendas» (Robinson, 2010: 3).

For present purposes I shall call this phenomenon *expectation-driven reductionism*. Its main source is the «incentivisation» of science. To be sure, this argument does not imply a general unease with reductive reasoning. On the contrary, reduction is needed in most scientific explanations. It is the process by which the scientist is isolating a domain of inquiry from the contingencies surrounding its manifestations so that its essential qualities may be fathomed. It lies at the heart of most scientific methods that questions and problems are limited, and that uncontrollable factors must be reduced or excluded. In this sense reductionism is a term used for the ambition to understand the nature of complex things by reducing them to the interactions of their parts, or to simpler or more fundamental entities. What I am opposing to in this article is strong multi-level reductionism, i.e. the ambition to reduce entire levels of scientific explanations to lower level accounts. While it is commonly held among philosophers that reductionist explanation is a very demanding task, scientists are often cautious when lending their support to multi-level reductionism in terms of currently established theories. However, as I argue below, it seems that scientists are far more optimistic when referring to the *future* advancements of science.

It is not unlikely that expectations and promises have always been essential in providing the dynamism and momentum upon which science depends. The ability of researchers to promise explanations of entire areas of human or social behaviour is increasingly valued in an institutional framework where the resources are scarce and the stakes are high. Being able to explain entities or phenomena that previously belonged to higher order disciplines, say, claims that neurolinguistics can replace comparative literature in explaining canonical texts, or that neuropsychology can replace political theory in explaining voter behaviour, has become an effective strategy when attracting funding and resources (Holtzman & Marteau, 2000; Levidov & Marris, 2001; Brown & Michael, 2003). But while pushing the promise of their findings or findings-to-be, the promises are often unbalanced and insufficient to ground sound scientific reasoning.

Again taking the long view of intellectual development, fetishising such priority in research caters to the superstition that the first route into a new field is the only or best route, and that the old routes are exhausted. Steve Fuller has carefully described this dynamics: «Of course, if enough people pay long enough lip-service to this superstition, it can turn into a self-fulfilling prophecy, at which point it becomes honoured as a "research tradition" dominated by rituals and patronage that are very hard to avoid or escape» (Fuller, 2006: 52).

Finally, a brief remark on the approach of this paper. The paper is situated within the emerging field of applied epistemology (Irzik, 2009) or «political epistemology» (Budtz Pedersen, 2011). Basically, this strategy consists in accepting the influence of institutional factors on scientific knowledge production while at the same time recognising that social aspects do not necessarily impede scientific progress but might – under the right conditions – promote it. The assumption is that institutional arrangements and incentives may favour the division of cognitive labour in ways that encourage scientific diversification and pluralism, while other arrangements may foster convergence and encourage researchers to pursue other methodologies, problems and experiments (cf. Kitcher, 1990). More particularly, the aim of political epistemology is to convey a better understanding of how the science system operates within the broader framework of social institutions. In this sense there is a clear commonality between applied epistemology and post-constructivist science studies, following the suggestion by Collin (2011: chap. 10).

2. Reductionism

In a recent journal, eleven leading neuroscientists published a manifesto that attracted much attention (Elger *et al.*, 2004). The authors announced that «within the foreseeable future, it will be possible to explain and predict psychological processes such as sensations, emotions, thoughts, and decisions on the basis of physiochemical processes in the brain». The neuroscientists expected the results of their research to lead to a profound revision in human self-understanding: «We stand at the threshold of seeing the image of ourselves considerably shaken in the foreseeable future» (Elger *et al.*, 2004, 37). Of course, it would be entirely welcome if neuroscientists were able to solve a problem that has generated centuries of philosophical controversy. Quarrels over areas of competence lead nowhere (Bennett and Hacker, 2003). But as evident in the quote the most recent contributions to this debate seem to have taken an unforeseen turn. The problem has actually put the focus on the future of all human and scientific cognition, and has thus pulled natural scientists into the domain of philosophical speculation.

By unpacking such a broad and fuzzy metaphysical agenda into more detail, we may come to a better understanding of the attractiveness of reductionism. First of all the emergence and growing acceptance within recent years of a new biological reductionism seems to have different motivations, including new technologies for studying the dynamics of brain. When we investigate the philosophical foundations of reductionism we need to ask what makes reductionism an attractive position as opposed to a moderate, pluralist or otherwise non-reductionist approach. Scientific reductionism has probably existed ever since the origin of modern science. Plausible reasons for this could be (i) that reductionism generates a clear programmatic statement within a discipline; (ii) that it enables convergence towards specific epistemic goals; or (iii) that it is based on beliefs of – real or anticipated – explanatory power and predictability (Rohrlich, 2001). Here, I want to suggest an additional reason (iv), which claims that incentives imposed from outside science play an equally important role in fuelling reductionist programmes. Taking into account that scientists in a number of situations have to extrapolate promises and expectations of future scientific breakthroughs, reductionism becomes an effective strategy of claiming the relative importance of a discipline above others.

Before proceeding any further, it is instructive to examine some simple definitions of reductionism. Reductionism can refer to:

(A) a methodological approach that explains the nature of complex things by reducing them to the interactions of their parts, or to more fundamental levels of explanation; or
(B) a distinct philosophical position which claims that a complex system is either understandable in terms of lower levels accounts (that the system is nothing but the sum of its parts) or higher levels accounts (that the parts are nothing but the reflection of the system).

Following (B), we can call *X*-reductionism the claim that *X* (e.g. religion, mind, intentionality, morality, norms) can be explained by reducing it to certain non-*X* causes, which then constitute the reductionist explanation. According to this typology reductionism is by no means restricted to explanatory programmes in the natural sciences. For example, sociological reductionism is known as the ambition to explain whole levels of human cognitive behaviour as mere subfields of the social. Marxist economists, for instance, explain political ideas as subordinated to economic laws, and constructivist sociology of science proclaims to explain all scientific claims as mere sub-phenomena to the social (Barnes, Bloor & Henry, 1996).

Reductionism does not exclude the existence of emergent properties, but it implies the ability to fully understand these properties in terms of the processes that compose them, or of which they are a lower level reflection. Hence, we can identify reductionism with the project of explaining the sciences of complex things, i.e. human or social sciences, by appeal to the laws that govern the behaviour of their constituents (Dupré, 2006).

Daniel Dennett has attempted to state why he thinks reductionism is appealing to scientists. He suggests that we search for a more coherent doctrine of reductionism that conveys a better understanding of the actual scientific practice. Dennett therefore supports a type of reductionism, which he claims is nothing more than merely materialism, by distinguishing between what he calls «crane reductionism» (CD) and «greedy reductionism» (GD). GD is the principle that every explanation in every field of science could be reduced to particle physics or string theory. This type of reductionism deserves the criticism imposed on reductionism in general, Dennett says, because the lowest-level explanation of a phenomenon, even if it exists, is not always the best way to explain it. In contrast, on Dennetts account, crane reductionism is a legitimate scientific enterprise that builds its understanding of reality by gradually accumulating sound explanations. Implied is a type of reductionism that unifies theories or disciplines only in the weak sense of requiring that they avoid contradicting one another (Dennett, 1995: 81).

Whereas crane reductionism means explaining a thing in terms of what it reduces to, for example, its parts and their interactions, greedy reductionism emerges when: «In their eagerness for a bargain, in their excitement to explain too much too fast», scientists and philosophers «underestimate the complexities, trying to skip whole layers or levels of theory in their rush to fasten everything securely and neatly to the foundation» (Dennett, 1995: 82)[2].

Setting aside for a moment Dennett's argument in favour of reductionism the limits of reductionism become evident at levels of organisation characterised by higher degrees of complexity, including phenomena such as culture, neural networks, ecosystems and other complex systems formed from assemblies of large numbers of interacting parts (Dupré, 2006). For example, claims that psychology is «based» on biology, or that sociology is «based» on economics, are often met with reservations. Such claims are difficult to substantiate empirically even though there are clear connections between the fields. For instance, most would agree that psychology is constrained by biology, and that convincing explanations of mental phenomena such as memory or perception need to take into account relevant clues from physiology and neuroscience.

In his book *Human Nature and the Limits of Science* (2001), John Dupré identifies a number of additional reasons for what motivates scientific reductionism. He defines reductionism as «the tendency to push a good scientific idea far beyond the domain in which it was originally introduced, and often far beyond the domain in which it can provide much illumination» (Dupré, 2001: 74). Dupré further contends that «devotees of these approaches are inclined to claim that they are in possession not just of one useful perspective, but of the key that will open doors to the understanding of ever wider areas of human action» (Dupré, 1994: 374).

Finally, it is worth noticing that the apparent attractiveness of reductionism stems from the fact that philosophers are generally keen to avoid the dualism urged upon them by the number of ontological primitives within the ontology. Ontological reductionism serves the function of reducing the number of ontological primitives. Philosophers welcome this because every ontological primitive demands a special explanation for its existence (Sterelny, 1998). For instance, if we maintain that «life» is not a physical property we must give a separate account of why some objects possess life and why others do not. This seems more often than not to be a difficult task.

---

[2] An example of greedy reductionism, named as such by Dennett, is the behaviourism of B. F. Skinner (Dennett, 1995: 395). Dennett writes: «Skinner proclaimed that one simple iteration of the fundamental Darwinian process – operant conditioning – could account for all mentality, all learning, not just in pigeons but in human beings. [...] Skinner was a greedy reductionist, trying to explain all design (and design power) in a single stroke» (Dennett, 1995: 395). However, Skinner himself understood his view as anti-reductionist. In *Beyond Freedom and Dignity* he wrote that while mental and neurological states exist, behaviour should not be explained with recourse to either.

Reductionism is thus one of the options available to those who think that humans and human minds are part of the physical world and should be explained only by references to physical entities. Hence, I take reductionism to be the project of integrating such mental notions as «thought», «intention», «reasons», and so on, with the natural world by showing them to be natural phenomena. The inspirations for this project are the more well-known reductions of science: of the heat of gases to molecular motion, of lightning to electric discharge, of the gene to DNA molecules, and so forth. Reductionists hope to show similar relations between mental states and neurophysiological activities in the human brain (Sterelny, 1998).

Here, I will not be able to rehearse all relevant arguments pro et contra reductionism. Instead I will highlight a simple but strong response prevalent in recent philosophical debates. According to this argument, the reduction of mental states to neural states encounters the problem of «the indispensable fact of the human observer». What this means is that neurobiological explanations of consciousness on most accounts accept that humans have the ability to form representations, only to consider these representations «a magical show that the agent lay on for himself» [3]. This concept of a consciousness as an epiphenomenal entity (or «show») is derived from a representational theory of the mind, i.e. the notion that things are present to the agent by virtue of being represented or modelled in the brain. However, as recently argued by Raymond Tallis (2011a; 2011b), «you cannot get to representation, without prior (conscious, first-order) presentation… so the latter cannot explain the former» [4]. In other words, it seems a to be shortcut to get rid of the «inner monitoring» of the agent's responses to her sensory stimulation by an abstract ontological commitment to reductionism. What neuroscience sees as a complex pattern of dynamic neural activity, is hardly the kind of thing that explains the uniqueness of agential consciousness.

The problem is that vague concepts such as the «internal observer», «inner monitoring», or the «magical show» (which are recurrent terms used by neuroscientists) exactly constitute the problem that these theories are bracketing. Tallis (2011a) shows that it is impossible to escape a certain privileged position. If we allow for a special kind of feedback loop (the first-order observer) we need to ask «who or what observes?», and a «privileged position by virtue of what?». How can there be a privileged position in the material world of which the brain is a part? As agents, we are convinced of the irreducible distinctiveness and causal effectiveness of our minds. Moreover, as scientifically informed persons who reflect on their own position in the natural world, we are at the same time justified in believing that the universe is unified and includes us as part of nature [5].

Confronted with this argument, reductionists will insist on a causally closed world, and pursue one of two strategies: (i) An eliminativist position that attempts to collapse

---

[3] Here, in the words of the evolutionary psychologist Nicholas Humphrey (2011).

[4] Tallis continues his arguments: «Neuroscientists of consciousness try to elude this obvious objection by asserting that representations are not (necessarily) conscious. In fact, all sorts of aspects of consciousness are not conscious after all». According to Tallis it is a trend among neuroscientists to think that already before human consciousness, animals were engaged in some kind of «inner» monitoring of their own responses to sensory stimulation. But, Tallis argues, «what is inner about unconscious processes, material events in the material brain? And how can they amount to monitoring? These questions are not silenced by the… reassurance that consciousness is the product of some kind of illusion chamber, a charade» (Tallis, 2011b).

[5] It is in virtue of this ontological commitment to monism that many philosophers are keen to avoid the dualism urged upon them by splitting the world into nature and mind. It is controversial whether ontological monism is compatible with «emergence», or if models of emergence are purely epistemic and exist only through the description of a system, and does not exist at a fundamental level. Here I will not pursue these questions any further.

the mind-brain relation by showing that the phenomenon of consciousness can be reconciled with the assumption that consciousness is nothing but an epiphenomenon. (ii) A non-eliminativist position that turns away from consciousness and towards the world. From the viewpoint of a mentalistic mind-body ontology, the latter position tries to justify the mind's phenomenal distinctiveness and causal effectiveness.

Again, the major stumbling block for solutions such as these is the methodological fiction of an «elusive view from nowhere, which relies on the problematic move of disengaging the objectivating perspective of the observer from the participant perspective of those who engage in scientific practice» (Habermas, 2007: 28). If, on the other hand, there is no way of getting around the obligatory complementarity of the observer's account relating to the scientific object (third person) and the participant's involvement in the scientific practice (first person), we need to make an epistemic turn (Bandura, 2008; McDowell, 2004)[6].

I conclude this section with the claim that because neural activity is a necessary condition of consciousness, it does not follow that the concept of consciousness should be recoined in neurological terms, still less that mental processes can be explained (exhaustively) by biological processes. Rather, it seems desirable to reconcile ontological monism with epistemic pluralism. For this to be a viable option, we should question the assumption that we substantially enrich our explanations of how the mind is functioning by reducing mental states to neurophysiological states. Of course, this argument does not reject the utility of neurally-based models for understanding significant psychological phenomena. However, it implies that research on non-neural cognitive psychology should enjoy equal priority. In the next section I elaborate this point further by showing its consequences for the use of models in science.

3. The strategy of model-based science

The philosophical responses to reductionism have been multivocal. There is a vast literature on this debate which the current analysis needs not address. For my purposes it does not matter how mental states relates to neurophysiological states as long as it is granted that different levels of explanation can coexist within a unified reality. While reality is admittedly ontologically homogeneous, it is diverse from the point of view of explanation[7]. In philosophy of science this strategy has especially been associated with the suggestion of a «model-based science» (Godfrey-Smith, 2006). On this account, the challenge for reductionism is that scientific models are only useful for a limited set of

---

[6] It goes without mentioning that the reducibility of the participant's perspective has been highly debated. Dennett, for instance, has leveled a radical skepticism towards human subjectivity with appeal to a forthright endorsement of evolutionary reductionism (Dennett, 1982; and further Dennett, 1991, 2002 and 2003). Even if this reflection on how scientific practice is semantically constituted from *within* the participant's perspective has some plausibility, it provides no easy recipe for the question of how human minds can understand itself as part of natural evolution.

[7] This position is reconcilable with ontological naturalism. Indeed, within a non-reductive and multi-scale ontology intentional and social explanations enjoy equal legitimacy to physical explanations. They represent natural facts in the sense that they are referenced by empirical sciences, but are divided into levels of explanation that are relatively autonomous – i.e. physical, biological, intentional or social. The rigidity of physical explanations might be higher than intentional or social explanations, but rigidity is only one criterion among others, as I shall argue in next section (for a further discussion, see Caro & Macarthur, 2006; List & Pettit, 2011).

questions. Dupré, already cited as an adherent of multi-level explanation, has given the following account of this strategy.

> «Evolutionary models, to take one example, have been enormously successful in helping us to understand the processes by which new forms of biological organisation emerged from earlier ones. […] But human behavior, though it is certainly an aspect of a biological entity that emerged from features of earlier biological entities, is very little illuminated by standard evolutionary models. Human development and human social organisation are two crucial aspects of the origination of human behavior that standard evolutionary models abstract away from» (Dupré, 2006: 5).

Dupré further argues that the attempt to make this extended abstraction viable is to provide an «imperialist» model of human behaviour:

> «It is no doubt a natural tendency, when one has a successful scientific model, to attempt to apply it to as many problems as possible. But it is also in the nature of models that these extended applications are dangerous. The abstractions that work well in one context may eliminate what is essential in another» (Dupré, 2006: 6).

If we take reductionism to imply that the best scientific strategy is to reduce explanations to the most fundamental entities, the reductionist thus maintains that the atomic explanation of a substance is always preferable to the chemical explanation, and that an explanation based on smaller parts (e.g. quarks) would be even more preferable, as outlined above. Reductionism in this sense is the position that all scientific explanations should be reduced to a single set of basic explanations, or, in the absence of a global explanatory reduction, that less fundamental sciences should be reduced to more fundamental sciences.

Before lending support to this type of reasoning consider a number of objections. First, every explanatory project is contingent on how we conceive our intellectual pursuits. It is possible that certain reductive explanations will open doors to the understanding of specific aspects of phenomena that we wish to explore in more depth – viz. the role of reduction as outlined in the introduction. However, the limit of global reductions becomes evident from the fact that the lowest-level explanation of a phenomenon, even if it exists, is not always satisfying to our cognitive interests. While it is commonly accepted that we can obtain important knowledge by understanding the relation of chemistry to physics, or the relation of microbiology to chemistry, similar statements become misleading or irrelevant when we consider less rigorously defined disciplines such as biology, psychology, sociology and the like.

A better way of approaching the difference between scientific disciplines is to understand scientific theories as «models» (Giere, 1999; Kitcher, 2001; Downes, 1992). Models are used by scientists to represent specific aspects of the world. These models include physical models or diagrammatic representations, but the models of most interest to the present argument are theoretical models. These are abstract objects whose structure bear a resemblance to «aspects» of objects and processes in the world. Scientists use models to gain an understanding of a complex real-world system by means of understanding a simpler, hypothetical system that resembles it in relevant respects (Godfrey-Smith, 2006: 726).

Any particular model have the function of true, but mediated statements about the world. Statements are true only in relation to the abstract model, e.g. in the same way explicit definitions are true. The empirical question – the question of realism – is how well the resulting model «fits» the targeted real-world system. No matter how precise, the fit is always partial and imperfect. There is no such thing as a perfect model. That, however, that does not prevent the models from providing the agent with deep and rigorous insight into the workings of the relevant scientific domain (Giere, 1999; Kitcher, 2001).

This philosophical notion of models, sometimes referred to as the semantic view, understands scientific theories as a family of models, which are used to represent real-world systems. A clearer account of this strategy results from comparing models to maps. Maps represent regions of the world from a particular point of view depending on the cognitive and pragmatic interests of the map-maker. Imagine, for instance, four different maps of London: (i) a street map, (ii) a subway map, (iii) a map of the electricity grid and (iv) a geological map. Each map represents the city from different perspectives. Which map we find to be the most significant depends on our cognitive and practical interest. For example, a driver, a subway commuter, or a construction-worker is likely to find relevant knowledge in each of the maps, and is able to navigate accordingly.

In this sense, maps (or models) demonstrate how a strong sense of truth and objectivity can be reconciled with significant levels of context-dependence and conventionalism. In particular, a map can provide representations of aspects in the world and yet function in accordance with conventions (e.g. about coastal lines, rivers, borders); epistemic interests (a map may focus on political or social aspects as opposed to physical aspects); and social contingencies (e.g. symbolic content, geographical orientation etc.) (Cordero, 2008: 101ff.).

What a map describes, what it includes and leaves out, its scale and accuracy, is determined by its intended use. No matter what it portrays a map will be useful only to the extent that the relevant information it provides is accurate within specified standards, even if the map is rigorously limited in this respect. If we add the intensions of the agent or community of agents, the following formula appears:

> «Agents (1) intend; (2) to use map, *M*; (3) to represent a part of the world, *W*; (4) for some purpose, *P*. So, agents specify which similarities are intended, and for what purpose» [8].

Maps have many of the representational qualities we need for understanding how scientists represent the world. There is no such thing as a perfect map. Nor does it make sense to question whether a map is true or false in the traditional sense. A map may, for example, be more or less accurate, more or less detailed. Nevertheless is it possible for a map to correspond with the target system in numerous ways. Some ways of constructing maps provide resources for capturing specific aspects of the world more or less well. But no single map will capture all aspects of reality (Giere, 1999: 79) [9].

While the notion of similarity may be intuitive, precision demands that the system is represented or measured in a regulated fashion, if evaluations of resemblance are to be attempted. Typically, measurements or data from the target system are arranged into a «data model», which provides the basic entities and relations to which the theoretical model is compared. In this case, accordance between data and prediction is a good indicator of the fit between model and world, or, more cautiously, a good indicator that the proposed model fits better with the world than alternative candidates (Giere, 1999: 75) [10].

---

[8] Here, I am following closely Giere (2010: 275). Notice that the formula is designed to eliminate the problem of multiple similarities. Giere calls this «the intentional conception of scientific representation». Admittedly this version presupposes a notion of representation, but, following Giere, we are not obliged to give a non-circular account of representation (Giere, 2010).

[9] Using the example of a person navigating the subway, Kitcher illustrates the possibility of generating and transmitting knowledge about parts of the world which are independent of observation: »Even if we were not around to watch the [map reader's] performance, she would still go through the same psychological states and perform the same actions« (Kitcher, 2001: 26). In this sense, the map metaphor is intended to capture a standard realist intuition while setting aside metaphysical presumptions (Sismondo & Chrisman, 2001: 41).

[10] Notice that the degree to which models are taken to represent the world is subject to controversy. What is frequently called an *instrumentalist* use of models amounts to treating models as calculating devices. More committed are idealized models, which can be seen as either false-but-handy simplifications

It is plausible that some things to which the model *does not* apply have causal connections to the things that it represents. Even if we try to shield off everything we know to be relevant, such shielding is rarely perfect. And we cannot shield off against every unknown causal factor. So, the only candidate model that might be a perfectly fitting model would have to be a model fitting everything perfectly. The prospects for constructing such a model are, however, very low (Giere, 2010: 273).

The argument outlined above has a number of consequences of how to conceive explanations. According to several recent contributions in philosophy of explanation, we are advised to make a distinction between the «significance» (or relevance) of an explanation and its «robustness». Explanations are part of an intentional practice in the sense that explanations are provided in an act of communication that is profoundly context-dependent. More precise, an explanation is developed as an answer to an explanation-seeking question (Faye, 2012). Which explanations we intend to pursue varies with respect to the context in which the question is asked, including the background assumptions and epistemic goals of the agent. Hence, there is no a priori theory of explanatory relevance, for instance, that explanations of the most fundamental entities are always preferable to explanations of higher order systems (e.g. mental, cultural, economic systems). As Kitcher convincingly has argued there is an inexhaustible number of correct (or approximate correct) explanations, but what science should aim at is delivering «significant» explanations (Kitcher, 2001: 65) [11].

This has several implications for the way we conceive reductionism and the prospects of applying a biologico-neurological understanding in e.g. economics, political science or the humanities. If we accept that there is no total or universal perspective, then all perspectives should be understood as partial representations of their objects. There is something real each perspective is a perspective of. In this respect, the humanities and social sciences are perfectly legitimate sciences that deliver partial models of human and social behaviour; their explanations might not be furnished by the same rigidity or generality as the natural sciences, but they deliver significant knowledge that accord with the cognitive interests in the epistemic community [12].

Neo-darwinians such as Daniel Dennett or Steven Pinker argue that people who are opposed to reductionism may be searching for a way of salvaging some sense of a higher purpose to life, in the form of some kind of non-material or super-natural entity. I disagree. Biological reductionism limits our understanding of complex organisations and leads to trivial and irrelevant claims that convey a general but unwarranted admiration for simplicity and unity. As scientific insights into genetics, ecology and other complex systems have revealed, if there exists «a systematic organisation of the truths about nature from which objective explanations may be drawn», science has generally been taking a step further away from it, not closer to it (Kitcher, 2001: 68-69).

In his book *Consilience: The Unity of Knowledge* (1998), E. O. Wilson passionately believes that we are progressing towards a synthesis of science, and that we will reach a

---

of the real-world system, or as approximately true representations of some of the forces operating in the system. For a further discussion, see Lloyd (1998).

[11] Obviously, to discuss in detail the relationship between explanation and context would be to go too far in this paper. But I think it can be safely said that the scope of explanations is flexible for the purpose of variable depth and epistemic interests.

[12] Again, the fact that e.g. humanities use «intentional» explanations limits their possibility of constructing highly integrative theories, or theories which are highly economical with respect to basic principles and laws. The role of intentional explanations is to a large extent dependent on the role ascribed to them by the actors themselves, and therefore cannot be substantially simplified (Collin, 2011, *forthcoming*).

point where we can unite the sciences and humanities. Yet, Wilson is disarmingly silent about the ultimate grounds for this belief: not evidence, but faith seems to be the basis of his reductionism. I will not try to recount the argument here. However, it is striking to observe that a large part of reductionist programmes are motivated not by adherence to evidence but by defending the view that science will eventually produce the explanatory power sufficient for a successful reduction. The problem is that reductionists such as Wilson never states exactly by which means we have ground for believing in the success of reductionism. He refers uncritically to anticipations of future scientific studies and uncritically accepts any lower-level explanation, even in cases where the lower-level explanation is of no relevant to understanding the target object.

Supporters of the neurobiological approach make a tempting but invalid inference from an ontological condition – that humans admittedly belong to the biological and physical world – to an explanatory conclusion, i.e., that humans and their cultural artefacts may be explained in biological and neurological terms. Such reductionist declarations neglect the fact there is vastly more relevant aspects to human behaviour than understanding the neurophysiological process that provides us with the fundamental capacities for our behaviour. Of course, this should not hollow the fact there are interesting questions concerning the neural capacities of human action that neurophysiological models should attempt to address. But it is in the nature of models that they are limited, and that the abstractions that work well in one domain may eliminate what is essential in others. In the next section I shall explore the relation of science and policy as a possible driver of reductionism.

4. The turn to science policy

In the remaining of this outline I want to make an untraditional turn. Having considered what reductionism implies and having argued in favour of a model-based understanding of science, I turn now to the institutional dynamics of science. More specifically, I shall suggest an alternative account of why reductionism continues to attract adherents and even seems to thrive well in the current scientific landscape. Arguably the configuration of research programmes is strongly tied to the epistemic interests in the community of inquirers and nourished by the plurality of academic interests and disciplines. Yet, science also connects with a much larger set of social interests. Society entertains research projects in all sorts of directions. Epistemic interests constantly compete with other interests prevailing in the broader societal context. Projects of great scientific relevance need to be negotiated with political priorities, particularly in the case of projects requiring long-term funding.

Whatever else science may be it is also an institutional process. Understanding the way in which institutional processes works is surely one fundamentally important approach to understanding science. There is a host of important questions about how the production of knowledge interacts with the rest of society that requires both institutional and philosophical reflection (Kitcher, 2001; Irzik, 2009; Douglas, 2009; Radder, 2010). In this respect there is room for what Kuhn regarded as 'nonscientific' values to enter into the scientific practice. These values might be personal, professional, social, political or a combination [13].

[13] Here, I am referring strictly to the proviso mentioned in the introduction, that this strategy consists in accepting the influence of social factors on scientific institutions while at the same time maintaining

Focusing on the diversity of epistemic agents, their perspectives, interests and values brings forward the policy framework of science. More particularly, I shall defend the thesis that reductionist programmes in current neuroscience are not simply driven by curiosity or cognitive interest but should be understood in conjunction with the incentives and arrangements derived from funding schemes and science policy [14]. In accordance with the reservations mentioned above, there are good reasons to doubt the value of multi-level reductionism. Besides generally underdetermined empirical claims, reductionist propositions rest for a large part on prospects of «future» scientific advances. This observation might be empirically trivial, but philosophically significant given the strong objection to reductionism. In the following I shall refer to this phenomenon as «expectation-driven reductionism».

Science is arguably becoming a more and more future-oriented enterprise in terms of expected utility and anticipated breakthroughs (Brown & Michael, 2003). Scientific knowledge is no longer simply about reviewing the established literature or making inference to the best explanations. The credibility ascribed to fashionable research programmes and the inferiority ascribed to «old» paradigms very much depend on the construction of promising «futures», viz. expected scientific results. In recent decades transformations in the societal role and position of science have ensured a more direct demand for high-impact knowledge, which has been theorised in concepts such as Mode-2 (Gibbons *et al.*, 1994), or Triple Helix (Etzkowitz & Leydesdorff, 2000). The new knowledge economy does not take the impact of science for granted; it has to be shown in applications and evaluations. This has resulted in an abundance of promises about future scientific breakthroughs such as new therapies, energy systems, or smarter materials, that are envisaged to provide new and better solutions for society (Felt *et al.*, 2007; Avadikyan *et al.*, 2003).

5. Expectation-driven reductionism

Funding and availability of resources in science is today highly competitive, and researchers have come up with a number of strategies of attracting the attention of reviewers, research councils, administrators and the media. From recent contributions in science policy studies we know that the enforcement of incentive structures, the commodification of research, and the struggle for resources are potentially creating a race to the bottom in form of promising technological and scientific results (Brown & Michael, 2003; Laudel, 2006; Horst, 2007).

Scientists are under constant pressure to explain their research priorities to policymakers and funding agencies, which have to take into account the broader societal perception of the pursued knowledge (van der Meulen, 2008). The ability to promise scientific and technological breakthroughs that radically depart from established wisdom or leads to radical innovations is increasingly valued in the conversion of resources and the attribution «relevance» in science (Felt *et al.*, 2009).

---

that social aspects do not impede scientific progress, but might – under the right conditions – promote it. One does not have to carry any constructivist luggage to appreciate the fact that science is influenced by extra-scientific norms and values.

[14] I take «curiosity-driven» research to be the standard definition of basic research, i.e. disinterested, truth-seeking science. This definition follows e.g. the suggestion of the European Research Council (ERC) which funds frontier research within all disciplines. Normally curiosity-driven research is contrasted with applied or mandated science which are driven by the demand of clients or societal challenges. For a further discussion of the morphology of research models and science policy doctrines, see Elzinga (2003).

The rapid growth of the science system, the rising costs of conducting science, and the financial constraints on government budgets all impact the funding of science (Radder, 2010). Many Western economies have undergone a change in the funding structure of their universities, with one of the major trends being a shift from recurrent block funding to performance-based funding and private grants. Shrinking university budgets cause researchers to rely more on external funding, which leads to a situation characterised by permanent competition and evaluation. This is intended insofar as the increased competition is assumed to lead to a higher degree of scientific excellence and enables funders to induce strategic priorities into science (Laudel, 2006: 489; Weingart, 2009: Budtz Pedersen, 2011).

However, these institutional arrangements also make researchers comply to the new conditions by altering their knowledge production. In an institutional framework where the policy goal is to stimulate higher quality (without changing the direction of science), these adaptations must be regarded as unintended side effects (Laudel, 2006: 489-490). For example, incentives in the way science is funded can affect the balance between applied and basic research, disciplinary and multi-disciplinary research, theoretical and technical research, etc. Or it can change the publication pattern of researchers, for instance, with regard to the lack of incentives to disseminate science publicly.

Second, the transformation in the knowledge producing system has been studied in relation to the effectiveness and fairness of peer-review in funding distribution. Interdisciplinary research is perceived as having fewer chances in peer-reviewed funding (Travis and Collins, 1991). Competitive funding has also been shown to affect the knowledge «content» by changing research questions and methodologies while forging the researchers to comply with the criteria set by the funding agency (Chubin & Hackett, 1990: 63; Wessely, 1998: 302) [15].

Moreover, from social studies of science we have learned that funding agencies and reviewers are highly sensitive to promises of future products and solutions, and that this plays a pervasive role when allocating money for research (Rip, 2009). Several authors have contributed to this literature on the «economy of expectations». Hessels *et al.* (2008) write that

> «[T]he criteria guiding the selection of research proposals often contradict the criteria guiding the selection of manuscripts for publication and the attribution of scientific recognition. Scientists are rewarded for making beautiful promises about the possible relevance of their research, not for realizing these promises» (Hessels *et al.,* 2008: 398).

In another influential paper, Borup *et al.* (2006) claims that:

> «By definition, contemporary science and technology is an intensely future-oriented business with an emphasis on the creation of new opportunities and capabilities. Novel technologies and fundamental changes in scientific principles do not substantively pre-exist themselves, except and only in terms of the imaginings, expectations and visions that shape their potential» (Borup *et al.,* 2006: 285).

A number of studies have examined the pervasive role of expectations in emerging technologies. For example:

- Membrane technology: Van Lente & Rip (1998).
- Neuroscience/ambient intelligence: Felt *et al.* (2007).
- Gene therapy/biotechnology: Horst (2007).

---

[15] In a comparative interview-based study of physicists working at Australian and German universities, connections between funding and adaptation strategies has been analyzed (Laudel, 2006). The study showed that strategies differ among scientists in the two countries due to their funding conditions.

- Pharmacogenomics: Hedgecoe & Martin (2003).
- Hydrogen storage: Van Lente & Bakker (2010).
- Fuel cells: Avadikvan *et al.* (2003).
- Xenotransplantation: Brown & Michael (2003).
- Stem cells: Brown & Michael (2003).
- E-commerce: Brown (2003).
- Nanotechnology: Nordmann (2004).
- Synthetic biology: Pauwels (2009).

While several of these studies have focused on the expectations of technical research (i.e. nanotechnology, biotechnology, synthetic biology), it seems reasonable to assume that an equally important feature of anticipating scientific results pertains to the explanatory expectations. Because scientists are constantly competing for resources at ever more diverse institutional levels they are encouraged to push their arguments to the limit, and sometimes of the limit, and exaggerate the prospects of future scientific explanations. As Rip explains, «a strategic game is being played… in which being first is more important than going in the right direction» (Rip, 2009: 669).

Having outlined these social and economic concerns regarding academic institutions, let us now return to the question of evolutionary-neurological reductionism and the escalating rhetoric of explaining human behaviour in neural terms. Consider the following examples in which the production of scientific «futures» play a significant role. Here is a passage from the famous neuroscientist Vilayanur Ramachandran at UCLA San Diego:

> «The discovery of mirror neurons in the frontal lobes of monkeys, and their potential relevance to human brain evolution... is the single most important "unreported" (or at least, unpublicized) story of the decade. I predict that mirror neurons will do for psychology what DNA did for biology: they will provide a unifying framework and help explain a host of mental abilities that have hitherto remained mysterious and inaccessible to experiments» (Ramachandran 2000).

Here, we can observe in detail how the success of neuroscience is depicted in terms of future explanatory success, thus constituting an example of expectation-driven reductionism. Mirror neurons are neurons in the monkey premotor cortex that are active both when a monkey produces an action such as grasping, and when it observes the action. No one yet knows quite why there is an overlap in patterns of neural activity. Ramachandran would like to find out, so he has made his priorities clear to the audience. We can review his record of publications and experiments, and conclude that research on mirror neurons is likely to yield significant knowledge in the future (for a further analysis of this example, see Gerrans, 2009).

To be sure, it is epistemologically legitimate to use future explanatory goals to highlight particular lines of research. But we should be very careful about the type of knowledge claims that we infer from such expectations, and make sure that any explanatory derivatives are secured against the prospects and are refuted if no solid explanations are given within a reasonable time. The intellectual significance of brain imaging techniques or interpretations of mirror neurons undergoes continuous revision as new studies are integrated with more complex models that reflect other aspects of human cognitive behaviour.

Consider now a second quote, this time from Jeffrey Krichmar, a cognitive neuroscientist. Krichmar writies that:

> «Work on artificially created machines that approach decisional-processes may ultimately obtain the parameters of human capacity, emphasizing that we should consider these potentialities beyond the mere utility of robots, for it may be that as a consequence

> of such advanced and iterative synthetic neural network development robots (and perhaps other computational devices), we will be able to manifest some form of consciousness. At first impression, this would allow a deeper insight to the mechanisms of consciousness, albeit not necessarily the same substrates as those of human or animal consciousness per se, but upon further reflection also instantiates moral consideration of what these conscious devices may feel» (Palchik, 2009: 2).

Even in the much cited paper by Francis Crick and Christof Koch, *Towards a neurobiological theory of consciousness* (1990), one finds examples of expectation-driven reductionism:

> «It is remarkable that most of the work in both cognitive science and the neurosciences makes no reference to consciousness, especially as many would regard consciousness as the major puzzle confronting the neural view of the mind and indeed at *the present time* it seems deeply mysterious to many... We suggest that the time is now ripe for an approach to the neural basis of consciousness. Moreover, we believe that the problem of consciousness can, *in the long run*, be solved only by explanations at the neural level» (Crick & Koch, 1990: 263, italics added).

Again we are presented with a fundamentally future-orientated explanatory programme. Relevant advances in neurosciences and artificial intelligence are displayed not in terms of current knowledge but in terms of future science. There may be other readings available. Be that as it may, the moral of these examples is that the claim of explanatory success is situated within an «anticipatory» epistemology which lacks the well-foundedness required in science. Expectation-driven reductionism describes the conditions under which it seems legitimate for members of a scientific collective to depart from established knowledge while not stating clearly the hypothetical character of their claims.

Features of this reasoning can be understood as an instance of what J. L. Austin called «performative utterances», or performatives. These are characterised by two features: (i) Performative sentences are not used to describe (or state a fact) and are not «true» when they succeed, and «false» when they fail; they have no truth-value. (ii) Second, to utter a performative sentence is not just to say something, but rather to perform an action. Austin calls this an illocutionary speech act. That is, a class of speech acts in which the speaker must indicate to the addressee that an act is performed (Austin speaks of «securing the uptake»), and which involves the production of «conventional consequences» such as commitments or obligations (Austin, 1975: 116f., 121, 139). Hence, in order to successfully perform a promise the agent must make clear to her audience that she is performing a promise and in the act of promising is undertaking an obligation. Searle's commentary on Austin's theory is even more precise. Searle distinguishes between «commissives», i.e. speech acts that commit a speaker to some future action (e.g. a promise), and «directives», i.e. speech acts that cause the hearer to take a particular action, following a request or advice.

In this sense expectation-driven reductionism is located at the intersection of promising scientific findings or findings-to-be and committing the recipient to provide the necessary framework for carrying out the research (i.e. funding, research infrastructure etc.).

A possible statement of this type of reductionism, which takes into account the performative directives just mentioned, is *that the lowest level explanation, E_1, of a given phenomenon, P, is more promising (will convey a better understanding) than the higher-level explanation, E_2, and that we should have a cognitive preference for the superseding theory T over the superseded theory S when allocating resources*. Dennett is implicitly relying on this strategy in his support for evolutionary psychology: «Darwin's dangerous idea [evolution by natural selection] is reductionism incarnate, *promising* to unite and explain almost everything in one magnificent vision. Its being the idea of an algorithmic process makes it all the more powerful, since the substrate neutrality it thereby possesses permits us to consider its application to just about anything» (Dennett, 1995: 82, italics added) [16].

The argument outlined above is meant to make sense of the rather imprecise explanatory ambitions that we often find in evolutionary-neurological approaches. Still, it is to a large extent unclear what role the production of scientific expectations has in scientific reasoning, and how such expectations connect to standard philosophical accounts of explanation. Every reliable explanation (reductive or otherwise) must rely on the methods and results of *established* science. It is ironic, therefore, that for numerous reductionist claims a primary objection is that they rest on anticipation of future science, and hence that the attraction of reductionism itself depends on extra-scientific foundations.

There a plenty of pitfalls when conducting scientific research. But we should always try to assess the plausibility assumption as carefully as possible before lending support to specific explanatory programmes. For instance, we can demand that proponents of excessively ambitious research programmes describe how to test the purported claims, or we can submit the anticipated explanations to standard rules of evidence. Given the lacking presentation of such evidence in most reductionist frameworks we are in no position to assume that the performance of the research will result in the anticipated closure. We are left with a stylised metaphysical theory in which the ideal symmetry between possible worlds is based on unattested truth-conditions. For it to be legitimate to use hypothetical statements in support of e.g. neuroscientific explanations it is not enough that they are useful in constituting a horizon for future research. Hypothetical claims should be based on reliable and well-founded methods before they are brought to bear on scientific priorities.

The basic position underlying this discomfort with future-oriented explanations is nothing but empiricism. Of course, as argued by Dupré (2006), one must be a sophisticated empiricist. The world does not simply speak to the scientist. The scientist must develop sophisticated instruments and models to understand the world. But if these models are not constantly evaluated and refined in interaction with empirical evidence they have little epistemic value. The objection, therefore, to expectation-driven reductionism is simply that it constitutes a divorce of scientific practice from empirical reality.

6. Conclusions

If the account offered above is on the right track it illustrates the enduring power of the metaphysics of science policy. Scientists have adopted a number of strategies to position themselves and their research that accord with the incentives urged upon them the by scientific institutions. These arrangements are most often designed to create more competition and higher research quality, but is documented to have widespread side-effects that potentially sidetrack the plurality of theories, methodologies and explanations. More fatally, I have claimed that science policy is one of several drivers of reductionism.

[16] To his defense, Dennett denounces any form of reductionism that embraces «mysteries or miracles at the outset». However, a few lines later, he writes that «it is only when overzealousness leads to falsification of the phenomena that we should condemn it» (Dennett, 1995: 82). In other words, Dennett has no conceptual tools for distinguishing between sound reductionism and expectation-driven reductionism at the level prior to initiating explanations.

Without trying here to sort out the notorious difficult questions of methodological reductionism, it seems plausible that, at least to some extent, the proliferation of reductionism is driven by expectations of future causal closure rather than established evidence. Of course, one cannot rule out reductionism a priori. But the burden of proof rests with the advocates of such imagined achievements. It is important to recognise that while neuroscience is becoming better and better to explain the causal structure of the brain it is by no means given that the results will provide ground for explaining such intricate intentional properties as human decision-making, ethics, aesthetics, political fairness, or synthesising these different perspectives into a single model. Instead it seems fruitful to accept scientific pluralism and the genesis of diverse scientific «models» each of which respond to our cognitive interests.

Two final remarks need to be said with regard to the division of labour between the humanities and natural sciences. First, it is important to recognise that reductionist programmes akin to those described above has flourished also in the human sciences. Marxism, Freudianism, structuralism, new criticism, social constructivism and so forth, constitute a raft of reductionist projects in the humanities. There has been a well-founded concern regarding the *exceptionalism* claimed for the humanities, rendering large parts of human intentional, cultural and linguistic behaviour untouched upon by the natural and cognitive sciences.

Secondly, even if today's scientific expectations do not deliver what they promise, they are already changing the research landscape. What is granted as putative explanations are likely to influence the choice of peer reviewers and funding agencies.

The consequence is that the disciplines and models that engage most vividly in anticipatory explanations will crowd out less promissory programmes. It is a plausible concern that this struggle for visibility and priority will impact the disciplines disproportionately, and that the humanities and social sciences stand to loose this race (Nowotny, 2007). There are signs that this danger is real. Evolutionary-neurological approaches present themselves as inherently hostile to the «intentional stance» associated with the humanities, and promise a better understanding of human beliefs, emotions and attitudes. This is unfortunate not only for science but for our culture, since the humanities and social sciences are crucial in providing the conceptual framework for political action and democratic decisions. While it is important to sustain a high level of funding for neuroscience this should not be at the expense of the humanities. Hence, interdisciplinary projects need to ensure a share of humanistic participation in order to stimulate more careful research questions, methodologies and proceedings (Budtz Pedersen, 2011).

If one accepts this reconstrual of the relationship between science and policy, a number of topics suggest themselves for further research. Expectation-driven reductionism is but one of many topics amenable to scrutiny in political epistemology. Addressing these issues will have to draw on a broad range of social, economic and political sciences as well as insights from ethicists, clinicians and philosophers.

Acknowledgements. This paper has been presented in various stages at the following workshops and conferences: *Humanities Between Constructivism and Biologism*, University of Copenhagen, 17 September 2010; *Danish Association for Philosophy*, University of Copenhagen, 5 March 2011; *Governing Futures*, Department of Social Studies of Science, University of Vienna, 24 September 2011; *What is Post-Truth Science?*, Centre for Semiotics, Aarhus University, 8 December 2011. I am indebted to the organisers and audiences for useful comments. Special thanks to Steve Fuller, Finn Collin and Jan Faye for suggestive remarks.

References

Ashcroft, R., & Martin, P. (2005), *What is Neuroethics? Report on Wellcome Trust Biomedical Ethics Summer School*, Oxford, 26-29 September 2005 (http://www.wellcome.ac.uk/).

Austin, J. L. (1975), *How to do Things with Words*, Second Edition, Oxford: Oxford University Press.

Avadikyan *et al.* (Cohendet and Heraud) (ed. 2003), *The Economic Dynamics of Fuel Cell Technologies*, Springer Verlag Berlin.

Bandura, A. (2008), «Reconstrual of free will from the agentic perspective of social cognitive theory». In: J. Baer, J. C. Kaufman & R. F. Baumeister (eds.), *Are we free?: Psychology and free will*, Oxford: Oxford University Press, 86-127.

Barnes, B.; Bloor, D., & Henry, J. (1996), *Scientific Knowledge A Sociological Analysis*, Athlone.

Bennett, M., & Hacker, P. M. S. (2003), *Philosophical foundations of neuroscience*, Blackwell.

Borup, M., *et al.* (Brown, N.; Konrad, K., & Lente, H. V.) (2006), «The sociology of expectations in science and technology», *Technology Analysis & Strategic Management*, 18/3: 285-298.

Brown, N., & Michael, M. (2003), «A Sociology of Expectations: Retrospecting Prospects and Prospecting Retrospects», *Technology Analysis & Strategic Management*, 15/1: pp. 3-18.

Budtz Pedersen, D. (2011), «A new start for the humanities is required for the 21st century», *Danish Yearbook of Philosophy*, Vol. 44, No. 1, forthcoming.

Burge, T. (2010), «A Real Science of Mind», *New York Times* (nytimes.com), The Opinion Pages, December 19, 2010.

Caro, M. D., & Macarthur, D. (2006), *Naturalism in Question,* Cambridge, MA: Harvard University Press.

Chubin, D., & Hackett, E. (1990), *Peerless Science: Peer Review and U.S. Science Policy*, Albany NY: State University of New York Press.

Collin, F. (2011), *Science Studies as Naturalized Philosophy*, Synthese Library, Dordrecht: Springer.

Cordero, A. (2008), «Pluralism, Scientific Values, and the Value of Science». In: Agazzi, E., & Minazzi, F. (eds.), *Science and Ethics: The Axiological Context of Science*, Brussels, Peter Lang: 101-116.

Crick, F., & Koch, C. (1990), «Toward a Neurobiological Theory of Consciousness», *Seminars in the Neurosciences*, 2:263-275.

Dennett, D. (1995), *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, NY: Touchtone.

Downes, S. (1992), «The importance of models in theorizing: a deflationary semantic view». In: Hull, D.; Forbes, M., & Okruhlik, K. (eds.), *PSA 1992,* Philosophy of Science Association: 142-153.

Dupré, J. (2001), *Human Nature and the Limits of Science,* Oxford: Oxford University Press.

— (2006), «The Disunity of Science». In: Newall, P. (ed.), *Galilean Library* (www.galilean-library.org). Retrieved September 1, 2010.

Elger, C. E., *et al.* (A. D. Friederichi, C. Koch, H. Luhmann, C. Malsburg, R. Menzel, H. Monyer, F. Rösler, G. Roth, H. Scheich & W. Singer) (2004), «The Manifesto», *Brain and Mind*, 6: 30-37.

Faye, J. (2007), «The pragmatic-rhetorical theory of explanation», *Rethinking Explanation*, Boston Studies in the Philosophy of Science, 252: 43-68.

— (2012), *After Postmodernism: A Naturalistic Reconstruction of the Humanities,* Palgrave.

Felt, U. (2007), *Taking European Knowledge Society Seriously*, European Commission: Brussels.

Gerrans, P. (2009), «Bubble trouble», *Times Higher Education* (www.timeshighereducation.co.uk/), 9 July 2009.

Giere, R. N. (1999), *Science Without Laws*, Chicago/London: The University of Chicago Press.

— (2010), «An agent-based conception of models and scientific representation», *Synthese*, 171: 269-281.

Godfrey-Smith, P. (2006), «The strategy of model-based science», *Biol Philos,* 21: 725-740.

Goldman, A. I. (1999), *Knowledge in a Social World*, Oxford: Oxford University Press.

Haalsten, H. (2007), «What to ask of an explanation-theory», *Rethinking Explanation, Boston Studies in the Philosophy of Science*, 252: 13-26.

Habermas, J. (2007), «The Language Game of Responsible Agency and the Problem of Free Will», *Philosophical Explorations*, 10/1: 13-50.

Hedgecoe, A., & Martin, P. (2003), «The Drugs Don't Work: Expectations and the Shaping of Pharmacogenetics», *Social Studies of Science*, 33/3: 327-364.

HOLTZMAN, N. A., & MARTEAU, T. M. (2000), «Will genetics revolutionize medicine?», *New England Journal of Medicine*, 343: 141-144.
HORST, M. (2007), «Public Expectations of Gene Therapy», *Science, Technology, & Human Values*, 32/2: 150-171.
HUMPHREY, N. (2011), *Soul Dust: The Magic of Consciousness,* Quercus Publishing Plc.
IRZIK, G. (2009), «Why Should Philosophers of Science Pay Attention to the Commercialization of Science?», *EPSA Epistemology and Methodology of Science*, 129-139.
KITCHER, P. (1990), «The Division of Cognitive Labor», *The Journal of Philosophy*, 87/1: 5-22.
— (2001), *Science, Truth, and Democracy*, Oxford: Oxford University Press.
LANDES, X.; ANDERSEN, M. M., & NIELSEN, M. E. J. (2011), «On Academia and Sea Elephants», *International Institute for Asian Studies,* 57: 41.
LAUDEL, G. (2006), «The art of getting funded: how scientists adapt to their funding conditions», *Science and Public Policy*, 33/7: 489-504.
LENTE, H. V., & BAKKER, S. (2010)), «Competing expectations: the case of hydrogen storage technologies», *Technology Analysis & Strategic Management*, 22(6): 693-709.
LENTE, H. V., & RIP, A. (1998), «The rise of membrane technology: from rhetorics to social reality», *Social Studies of Science*, 28(2): 221-254.
LEVIDOW, L., & MARRIS, C. (2001), «Science and Governance in Europe: lessons from the case of agbiotech», *Science and Public Policy*, 28(5): 345-60.
MCDOWELL, J. (2004), «Naturalism in the Philosophy of Mind». In: CARO, M. D., & MACARTHUR, D. (eds.), *Naturalism in Question*, Cambridge, MA: Harvard University Press, 91-105.
NORDMANN, A. (2004), *Converging Technologies. Shaping the Future of European Societies*, Luxembourg: European Commission.
PALCHIK, G. (2009), «Conference Report: Neuroscience & the Nature of Being», *Philosophy, Ethics, and Humanities in Medicine*, 4:9.
PAUWELS, E. (2009), *The Value of Science and Technology Studies (STS) to Sustainability Research in Europe*, unpublished.
PSILLOS, S. (2007), «Causal Explanation and Manipulation», *Rethinking Explanation, Boston Studies in the Philosophy of Science*, 252: 97-112.
PUTNAM, H. (2002), *The Collapse of the Fact/Value Dichotomy and Other Essays*, Cambridge, MA: Harvard University Press.
RACINE, E.; BAR-ILAN, O., & ILLES, J. (2006), «Brain Imaging: A Decade of Coverage in the Print Media», *Science Communication*, 28 (1): 122-143.
RADDER, R. (ed.) (2010), *The Commodification of Academic Research. Science and the modern university*, Pittsburgh, PA: University of Pittsburgh Press.
RAMACHANDRAN, V. S. (2001), «Mirror Neurons and Imitation Learning as the Driving Force Behind The Great Leap Forward in Human Evolution», *Edge.com*.
RIP, A. (2009), «Futures of ELSA», *EMBO Reports*, 10(7): 666-670.
ROBINSON, M. (2010), *The Privatization of Neuroscience: The University, The State and the Moral Aims of Science*, Somatosphere.net, Wednesday, November 10, 2010 (retrieved 11 January 2011).
ROHRLICH, F. (2001), «Cognitive Scientific Realism», *Philosophy of Science*, 68(2): 185-202.
SCHNEIDER, T. (2010), *Neurosociety. What is it with the brain these days?*, Retrieved 16 April 2011 from www.oxbionet.medsci.ox.ac.uk.
SISMONDO, S., & CHRISMAN, N. (2000), «Deflationary Metaphysics and the Natures of Maps», *Philosophy of Science*, 68(3): S38-S49.
STERELNY (1998), «Reductionism in the philosophy of mind». In: CRAIG, E. (ed.), *Routledge Encyclopedia of Philosophy*, London: Routledge.
STILGOE, J. (2011), «A question of intent», *Nature Climate Change,* 1(1): 325-26.
TALLIS, R. (2011a), *Apeing Mankind,* Durham: Acumen Publishing.
— (2011b), «A mind of one's own», The New Statesman, Published 24 February 2011 (www.newstatesman.com).
TRAVIS, G. D. L., & COLLINS, H. M. (1991): «New light on old boys: cognitive and institutional particularism in the peer review system», *Science, Technology and Human Values*, 16(3): 322-341.
VAN DER MEULEN, B. (2008), «Interfering governance and emerging centres of control». In: WHITLEY, R., & GLÄSER, J. (eds.), *The Changing Governance of the Sciences: The Advent of Research Evaluation Systems*, Dordrecht: Springer: 191-204.

WEINGART, P. (2009), «On Best Practice Rules of Publishing and Their Erosion», *Minerva*, 47:237-239.
WESSELY, SIMON (1998): «Peer review of grant applications: what do we know?», *Lancet*, 352(9124): 301-305.
WILSON, E. O. (1998), *Consilience: The Unity of Science*, New York: Knopf Publishers.

DAVID BUDTZ PEDERSEN

Department of Philosophy
University of Copenhagen, Denmark
davidp@hum.ku.dk