



Naturalisms in Philosophy of Mind

Steven Horst*

Wesleyan University

Abstract

Most contemporary philosophers of mind claim to be in search of a ‘naturalistic’ theory. However, when we look more closely, we find that there are a number of different and even conflicting ideas of what would count as a ‘naturalization’ of the mind. This article attempts to show what various naturalistic philosophies of mind have in common, and also how they differ from one another. Additionally, it explores the differences between naturalistic philosophies of mind and naturalisms found in ethics, epistemology, and philosophy of science. Section 1 introduces a distinction between two types of project that have been styled ‘naturalistic’, which I call *philosophical naturalism* and *empirical naturalism*. Sections 2 to 6 canvass different strands of philosophical naturalism concerning the mind, followed by a much briefer discussion of attempts to provide empirical naturalizations of the mind in Section 7. Section 8 concludes the essay with a consideration of the relations between philosophical and empirical naturalism in philosophy of mind, arguing that at least some types of philosophical naturalism are incompatible with empirical naturalism.

A casual observer of recent philosophy of mind would likely come to the conclusion that, amidst all of the disagreements between specialists in this field, there is at least one thing that stands as more or less a consensus view: the commitment to a *naturalistic* philosophy of mind. Almost everyone writing in philosophy of mind over the past several decades has described his or her theory as ‘naturalistic’. This includes the proponents of quite a wide variety of views: reductionist, eliminativist, informational, non-reductive physicalist, functionalist, computational, and evolutionary. Even David Chalmers, perhaps the most influential figure in the revival of property dualism in the 1990s, describes his position as ‘naturalistic’ (cf. Chalmers). At first glance, then, philosophers of mind might seem to have found at least one happy point of agreement at the turn of the millennium.

However, the very fact that attempts to ‘naturalize’ the mind have become something of a cottage industry indicates that there are features of the mind that at least *seem* to be very different from those that are paradigmatically ‘natural’. No one, after all, feels a need to try to ‘naturalize’ geology or chemistry. The phenomena treated by those sciences are

already regarded as 'natural', and their explanations are not seen as being importantly different from explanations of other natural phenomena.

But the mind presents us with several features that seem to be importantly different from anything we find in the physical or biological sciences. Many contemporary writers think that *qualia* (the subjective feelings of experiential states) and consciousness (subjective awareness) pose special 'hard problems' for the naturalist (Jackson, 'Epiphenomenal Qualia'; Levine; T. Nagel; Chalmers). To this list others would add intentional states, such as occurrent beliefs and desires, either on the grounds that they involve consciousness (Searle; Horst, *Symbols, Computation and Intentionality*; Siewert; Horgan and Tienson) or because they involve normative properties that cannot be derived from facts (Brandom, *Making it Explicit*; 'Modality, Normativity, and Intentionality'; Putnam, 'Why Reason Can't be Naturalized'). Some of these writers indeed see these features of the mind as posing deep problems for naturalism; but others join a much larger group of philosophers of mind in holding that these properties either *are already* 'natural' properties or else *can be explained by* more paradigmatically natural phenomena, particularly physical or neural phenomena. Most philosophers of mind tend to see qualia, consciousness, intentionality, and normativity as presenting *problems*, but not *insurmountable* problems, for the naturalist.

But things are not quite so simple. And the fact that they are not so simple ought to be foreshadowed by the very variety of views that can be styled as 'naturalistic'. If a reductionist, an evolutionary theorist and a dualist can each apply the label 'naturalist' to himself, it is very likely to prove the case, either that they are using the word in subtly different ways, or else that the word has become so bland and ecumenical as to be essentially useless.

While this is an important realization, it is by no means a new one. The ambiguity of the word 'naturalism' has been widely noted, and has been remarked upon for perhaps half a century now. The philosopher of science Ernest Nagel, in his 1955 presidential address to the American Philosophical Association, noted that 'the number of distinguishable doctrines for which the word "naturalism" has been a counter in the history of thought is notorious' (3). In their introduction to the anthology *Naturalism: A Critical Appraisal*, Steven Wagner and Richard Warner express a similar view:

Participants in current discussions of naturalism seem to assume that the meaning of 'naturalism' ('naturalist program', etc.), its motivations and – often – its correctness, one way or the other, are almost obvious. The historical situation makes such assumptions exceedingly unlikely. Philosophers have taken just about every possible stance with some manner of justification, and all of the main programs within this area ('naturalism', 'phenomenology', 'analytic philosophy', and so forth) have been open to sharp differences of interpretation by their adherents. (3)

In a similar vein, David Papineau begins his book *Philosophical Naturalism* with the question,

What is philosophical ‘naturalism’? The term is a familiar one nowadays, but there is little consensus on its meaning. . . . I suspect that the main reason for the terminological unclarity is that nearly everybody nowadays wants to be a ‘naturalist’, but the aspirants to the term nevertheless disagree widely on substantial questions of philosophical doctrine. (1)

Some philosophers, like Jesse Hobbs, have taken Papineau’s point that ‘nearly everybody wants to be a “naturalist”’ even further, raising the question of whether the word ‘naturalism’ is simply ‘a contemporary shibboleth’. Someone who came to this conclusion would, I think, be *half* right. The word ‘naturalism’ *does* tend to function as a kind of shibboleth – that is, as a word whose use distinguishes ‘members of the tribe’ from outsiders. And it is true that naturalism has become a kind of ideology in philosophical circles – that is, it is a widely shared commitment to a way of believing, speaking and acting whose basic assumptions are seldom examined or argued for. However, I think that this is not the whole story. The word ‘naturalism’ may serve as a shibboleth, but it is not *merely* a shibboleth. There may be a pervasive naturalistic ideology, but it is possible to articulate and examine some of its basic underpinnings. And if there is not a *single* view called ‘naturalism’ shared by the majority of contemporary philosophers of mind, there is nevertheless a way of bringing some order to the various views thus denominated, highlighting their commonalities as well as their differences.

I shall begin (Section 1) by making a distinction between two types of project that have been styled ‘naturalistic’, which I shall call *philosophical naturalism* and *empirical naturalism*. I shall then canvass different strands of philosophical naturalism concerning the mind (Sections 2–6), followed by a much briefer discussion of attempts to provide empirical naturalizations of the mind (Section 7). The essay will conclude (Section 8) with a consideration of the relations between philosophical and empirical naturalism in philosophy of mind, arguing that at least some types of philosophical naturalism are incompatible with empirical naturalism.

1. Two ‘Naturalistic’ Projects

The term ‘naturalism’ is associated with two distinct theses. The first of these is, in quite general terms, a *philosophical* position. In this first sense, a naturalist about X (e.g., consciousness, intentionality, knowledge, goodness) is someone who holds that X either *is* or *must be* nothing more than a natural phenomenon. This philosophical thesis, in turn, involves two types of commitments. One is to some sort of linkage between X and the phenomena that make up the domain of natural sciences like physics and neuroscience. This might be a commitment to a type of *explanation*. (For

example, that mental states are *reducible* to neural states or *explainable* as results of natural selection.) Or it might be a commitment to a type of *metaphysical relationship*. (For example, that mental states are *metaphysically supervenient* upon brain states or physical states.) The second commitment is to a rejection of entities and properties that are paradigmatically *supernatural*, especially immaterial souls. I shall call proponents of these views *philosophical naturalists*.

The second type of naturalistic thesis is more of a *methodological commitment*, concerning the relation between philosophy and the sciences. This is the thesis that philosophy does not have its own distinctive claims to knowledge, but rather is – or at least should be – continuous with the sciences, and is ultimately beholden to them as the final arbiters of the truth of any claims the philosopher might advance or scrutinize. Philosophy has important roles to play in commenting upon the sciences, in synthesizing the results of scientific inquiry, and as a kind of proto-scientific inquiry pursued in the hopes that a mature science will emerge. But it has no special role as a referee of scientific claims or methods, and can supply nothing of its own in the form of *a priori* reasoning. As Michael Devitt and Kim Sterelny put it:

Briefly, and roughly, we can divide philosophy's role in three.

(1) Philosophy's most basic task is to reflect upon, and integrate, the results of investigations in the particular sciences to form a coherent overall view of the universe and our place in it.

(2) Philosophy is concerned with certain problems in particular sciences, for example, in physics, biology, psychology, and mathematics. These problems arise in the most speculative and conceptually difficult parts of the sciences.

(3) Some sciences, or areas of sciences, are traditionally done in philosophy, in some cases, but certainly not all, because they are not mature enough to go out on their own: epistemology, logic, morals, politics and aesthetics. (We confess to having only the dimmest ideas about how to accommodate some of these within our naturalistic viewpoint.) (276)

I shall refer to this view as *empirical naturalism*, on the grounds that it identifies philosophical inquiry as either a form of empirical inquiry, or else a commentary upon it. We may summarize the commitments of philosophical and empirical naturalism as follows.

Philosophical Naturalism about X involves . . .

- Commitment to theses about relations between X and the objects, properties and laws of the natural sciences, particularly physics:
 - Types of *explanation* of X in unproblematically natural terms (e.g., reduction of X to physics).
 - Types of *metaphysical relationship* between X and the objects of the natural sciences (e.g., metaphysical supervenience of X upon physical phenomena).
- Rejection of paradigmatically *supernatural* entities and properties (e.g., Cartesian souls).

Empirical Naturalism about X involves

- Rejection of special *aprioristic* philosophical methods of knowing about X.
- Commitment to viewing one or more sciences as being the ultimate sources and arbiters of knowledge about X.

1.1. TERMINOLOGY

This distinction is not original, and has been noted by a number of other writers, albeit in different terms. Lawrence Shapiro, for example, distinguishes between a metaphysically oriented ‘Lego naturalism’ and a ‘methodological naturalism’. Papineau, in his entry on ‘naturalism’ for the *Stanford Encyclopedia of Philosophy*, makes the distinction in terms of ‘ontological’ and ‘methodological naturalism’. And Penco, Beaney and Vignolo identify

two basic strands within naturalism: one has an ontological orientation and the other has an epistemological one. From the ontological standpoint, naturalism implies the rejection of whatever supersedes the natural. From the epistemological standpoint, naturalism implies the rejection of the view that conceptual analysis is a source of *a priori* and unrevisable knowledge. (xxv)

These different sets of terminology are aimed more or less at the same distinction. I prefer to make the distinction in terms of ‘philosophical’ versus ‘empirical naturalism’ (as opposed to ‘ontological’ and ‘methodological’ naturalism) on the grounds that what I am calling ‘philosophical’ naturalism often involves commitments to particular types of methodology and explanation as well as to a metaphysical position such as physicalism, while those that I call ‘empirical’ naturalists often take an ontological position as well. (Namely, that the inventory of the universe consists in the theoretical posits of ideally completed sciences, whatever they might turn out to be.) I believe, however, that this is largely a difference in terminology rather than a substantive disagreement.

Of course, many philosophical naturalists are empirical naturalists as well, and vice versa. However, the two projects are defined by distinct sets of commitments; and at in at least some cases, these commitments may pull in opposite directions. For example, some philosophical naturalists employ exactly the sort of *a priori* philosophical reasoning that is rejected by empirical naturalists. And, more obviously, some philosophical naturalists pursue their arguments at a purely philosophical level, with little reference to the sciences, while some empirical naturalists are not interested in traditional philosophical problems like the metaphysics of the mind. I shall return, at the end of this essay, to the relations between philosophical and empirical naturalism.

1.2. NATURALISMS IN SPECIFIC AREAS OF PHILOSOPHY

In some areas of philosophy, such as epistemology and philosophy of science, the word ‘naturalism’ usually denotes *empirical* naturalism. Naturalistic

epistemology traces its roots to Quine's 'Epistemology Naturalized' (1969), a seminal statement of the view that philosophy is (or ought to be) continuous with the sciences, and of the rejection of *a priori* methods. Naturalistic philosophy of science represents a similar turn away from the *aprioristic* approach to science favored by the Logical Positivists and Empiricists, and towards pursuing a philosophical understanding of science guided by careful attention to case studies in the various sciences. (See discussions in Callebaut for a useful overview of the emergence of naturalistic philosophies of science.)

In ethics, by contrast, 'naturalism' usually denotes *philosophical* naturalism. G. E. Moore and R. M. Hare defined 'naturalism' in ethics as the view that ethical *terms* like 'good' can be *analyzed* or *defined* in non-ethical terms. Subsequently, as philosophy turned away from linguistic analysis, ethical naturalism has come to be understood as the thesis that ethical *properties* (like goodness) *supervene* upon non-ethical properties. In recent decades, however, there have also been attempts to provide empirically naturalistic accounts of ethical phenomena like normative evaluation and the moral emotions, particularly in evolutionary terms (cf. Richards; Ruse, *Taking Darwin Seriously*; Gibbard; Collier and Stingl).

Philosophy of mind is a more complicated matter. There are long-standing projects in philosophy of mind that are philosophically naturalistic, going back at least to Hobbes's attempts to reduce mental phenomena to motions in the body, understood mechanistically, and Descartes's counter-arguments that certain features of the mind (reason and language) cannot be explained mechanistically, and require the postulation of an immaterial soul. Such debates have continued to the present, though the field of contenders has been expanded to include such positions as non-reductive physicalism and eliminativism, and such non-reductive explanatory strategies as accommodation under laws and evolutionary explanation. In recent decades, however, philosophy of mind has also seen a movement to understand the mind in *empirically* naturalistic terms, by appealing to emerging work in the burgeoning fields of the cognitive and life sciences. While there is some cross-over between these two types of naturalizing projects, I shall discuss them separately before considering their relations to one another.

2. *Philosophical Naturalisms in Philosophy of Mind*

I suggested in the previous section that philosophical naturalism is centered around two commitments. The first is a commitment to accommodating our understanding of the mind to the natural sciences. The second is a rejection of such paradigmatically *supernatural* entities as Cartesian souls. A reading of the *Oxford English Dictionary's* entries on 'naturalism' and 'naturalist' will reveal that these words began to acquire both of these features – an association with the sciences and an opposition to the supernatural – in the sixteenth and seventeenth centuries. And indeed, it was around

this time, as well, that the word ‘nature’ (sometimes capitalized) came to stand for the universe as a whole. This should come as no surprise, as this was precisely the period in which what we now call the natural sciences (which would then have been called ‘natural philosophy’) began to be developed in Europe, and in particular the idea that there might be a *single* science, equally applicable to all observable phenomena. Before the seventeenth century, the prevailing assumption had been that the range of phenomena that could be explained ‘mechanically’ was rather narrow, excluding not only the motions of the heavens but also all organic processes. Descartes’s suggestion that organic processes were mechanical, and Newton’s synthesis of terrestrial and celestial mechanics, made the prospect of a unified science of nature seem to be within reach, raising the urgent question of whether the emerging mechanical view of the universe could be extended to encompass other areas of human understanding as well: the mind or soul, ethics, politics, and even theology.¹

We may codify a first attempt at characterizing philosophical naturalism about the mind in terms of a schematic definition:

General Schema: Philosophical naturalism in philosophy of mind is the view that all mental phenomena are to be accommodated within the framework of nature as it is understood by the natural sciences.

But this General Schema alone is not quite enough to exclude all views that would generally be considered non-naturalistic. Some philosophers, for example, equate ‘the natural’ with ‘the causal’ – i.e., with everything that enters into causal relations. On such a view, a God, an angel, or an immaterial soul that caused events in the natural world, or even outside of it, would count as a ‘natural’ entity. This, however, is inconsistent with the strand of naturalism that contrasts ‘natural’ entities, properties and processes with ‘supernatural’ ones. One might also think that Idealism is paradigmatically non-naturalistic, even in the case of Idealisms like those of Berkeley or Leibniz, which might be seen as treating the natural world as comprehensive (or very nearly so), but then interpret that world in terms of something more basic than the material bodies and physical properties, such as the coordinated perceptions of Monads. We thus need to supplement the General Schema with a caveat:

Caveat: A view cannot count as philosophically naturalistic if it (a) allows the existence of paradigmatically non-natural *entities* such as God, angels or Cartesian souls or non-natural *properties*, or (b) treats the world of nature as understood within the sciences as non-fundamental.

3. *Three Dimensions of Ambiguity*

Philosophical Naturalism, as characterized by the General Schema and the Caveat, is not really anything so exact as a shared *theory*. Instead, it is something on the order of a *theory-schema*. It is *only* a schema for theories

because there are several elements of this characterization that are ambiguous, and which different self-styled ‘naturalizers’ would fill out in different ways. There are, at very least, three axes along which this schema is ambiguous, which can be used to distinguish different varieties of naturalism.

1. Whether the claim about ‘accommodation’ is a claim about *explanation* or a claim about *metaphysical determination*, or both;
2. How we are to understand the phrase ‘the framework of nature as it is understood by the natural sciences’, and;
3. Whether the general schema is understood as a *positive* claim (that the mind *can* be so accommodated) or as a *normative* claim (that it *must* be so accommodated, or else some dire consequences follow).

Let us consider these issues in order.

3.1. EXPLANATION AND METAPHYSICAL DETERMINATION

Examinations of naturalism in philosophy of mind often mix together discussions of whether features of the mind such as consciousness and meaning can be *explained* by the natural sciences with discussions of *metaphysical* questions (such as whether mental states *supervene* upon brain states). For many naturalists, both sort of questions are deemed to be of great importance. And there *are* styles of explanation that are closely linked to particular types of metaphysical determination. For example, if everything about the mind can be explained, without remainder, by properties of the brain, then any two organisms with identical brain states will necessarily have identical mental states as well. However, *metaphysical questions* and *questions about explanation* are separable from one another: On the one hand, there are forms of explanation (e.g., probabilistic explanation and Humean generalizations) that have no metaphysical consequences. On the other hand, it might be the case that there are metaphysical necessities that are *epistemically opaque* (that is, beyond our ability to comprehend), and which consequently have no attendant forms of explanation to go along with them. This is most readily evident in the growing popularity, over the past two decades, of non-reductive and Mysterian views of the mind-body or mind-brain relation (McGinn; Pinker).

Reduction is an especially strong form of explanation, in which the features of the reduced system can be reconstructed as derived or constructed from those of the reducing system in the form of an axiomatic system. Such a view of inter-domain relations in the sciences was popularized by Carnap (*Aufbau*) and Ernest Nagel (*Structure of Science*), and was a mainstay of Logical Empiricist Philosophy of science, and reductionist philosophies of mind in the latter part of the 20th century were influenced heavily by this paradigm. However, several factors in the 1980s and 1990s led many philosophers to suspect that mind-body reductions might not be forthcoming. Rather than embrace dualism, however, many of them instead opted for

the view that mind-body relations are metaphysically necessary yet abidingly epistemically opaque, perhaps due to some special inability of the mind to understand its relation to the processes from which it emerges (McGinn; Pinker). On the other hand, other philosophers have been inclined to accept weaker, non-reductive explanations of mental phenomena, such as mechanistic² explanations (Bechtel and Abrahamsen; Wright and Bechtel), which do not entail metaphysical supervenience. So in examining a particular naturalistic claim it will be important to identify whether it is a claim about explanation or a claim about metaphysics or both.

3.2. THE FRAMEWORK OF NATURE AS UNDERSTOOD BY MODERN SCIENCE

Likewise, even once we have pinned down what we mean by ‘accommodating’ the mind within nature, the expression ‘the framework of nature as it is understood by the natural sciences’ is still rather vague. Just what our naturalistic schema *means* will depend heavily upon what one deems to be central to how the natural sciences operate, and how they represent the natural world. That is, it will depend upon what particular views one takes in philosophy of science on issues like the nature of explanation and the metaphysical commitments of the sciences. And this is a serious complication, because there are many alternative views on these subjects, as we shall see in Sections 4 to 6.

3.3. POSITIVE AND NORMATIVE CLAIMS FOR NATURALISM

And there is also a third axis of ambiguity: sometimes naturalistic claims are put forward as a kind of *positive* claim – a claim about how things are. These are a sort of second-order empirical prediction about how it will turn out in the long run. Positive empirical claims can often be put to the test and be shown to be true or false: it might turn out that some feature of the mind, such as consciousness, *can* be naturalized, or it might turn out that it *cannot*. This seems to be the tenor, for example, of Oppenheim and Putnam’s ‘Unity of Science as a Working Hypothesis’ (1958), as they write in a hypothetical tone that

It is not absurd to suppose that psychological laws may eventually be explained in terms of the behavior of individual neurons in the brain; that the behavior of individual cells – including neurons – may eventually be explained in terms of their biochemical constitution; and that the behavior of molecules – including the macromolecules that make up living cells – may eventually be explained in terms of atomic physics. If this is achieved, then psychological laws will have, in *principle*, been reduced to laws of atomic physics. (7)

But some naturalists have a dangerous tendency to slide into a different sort of claim that is *not* empirical or positive, but *normative*. They claim, in essence, that the mind *must* be naturalized, or else something unseemly

follows: that psychology cannot be scientific unless its objects can be explained in terms of something more fundamental, or that mental states do not exist unless they supervene upon physical states. Stich and Lawrence put the point as follows:

In recent years, many philosophers have put a very high priority on providing a 'naturalistic' account of intentional categories. Moreover, there is an unmistakable tone of urgency in much of this literature. Naturalizing the intentional isn't just an interesting project, it is vitally important. *Something dreadful* will follow if it doesn't succeed. And for many writers, we suspect, that dreadful consequence is intentional realism. (161)

We can see this same line of reasoning at work in Jerry Fodor's *Psychosemantics*. There, Fodor describes his project in the following terms:

Here, then are the ground rules. I want a *naturalized* theory of meaning; a theory that articulates, in nonsemantic and nonintentional terms, sufficient conditions for one bit of the world to *be about* (to express, represent, or be true of) another bit. (98)

Why is such a project important? In order to preserve the ontological credentials of intentional states:

The deepest motivation for intentional irrealism derives . . . from a certain ontological intuition: that there is no place for intentional categories in a physicalistic view of the world; that the intentional can't be *naturalized*. (97)

Fodor elaborates on this concern in *A Theory of Content*:

If it isn't literally true that my wanting is causally responsible for my reaching, and my itching is causally responsible for my scratching, and my believing is causally responsible for my saying . . . , if none of that is literally true, then practically everything I believe about anything is false and it's the end of the world. (156)

In the end, Fodor's motivation is normative: Intentional states like beliefs and desires *must* be real, and in order to vindicate their reality, they must be naturalized. Dretske puts the matter perhaps even more starkly, opining that, without a naturalized theory of content, we might need to 'relinquish a conception of ourselves as human agents' (*Explaining Behavior x*).

Likewise, seminal statements of eliminativism, like those of P. M. Churchland and Stich (*From Folk Psychology to Cognitive Science*), argue that intentional states are *not* ontologically respectable on the grounds that the sciences of cognition seem to be coming up with categorizations that do not map smoothly onto the intentional categories of belief and desire. The latter are thus 'unnecessary theoretical posits' that should be dispensed with, as we did with phlogiston. Again, the direction of argumentation is normative: *if* mental categories are to be ontologically respectable, they must prove their *bona fides* by being explained in terms of the natural sciences of the mind. And if they cannot do so, we must learn to live, or

at least to theorize, without them. On this point, Fodor, Stich, and Churchland were agreed (though Stich ('What is a Theory of Mental Representation?') has later repudiated this position): mental categories *must* be naturalized if they are to be ontologically respectable. And hence they agreed that there is a *forced choice* between some form of naturalization and elimination.

3.4. DISAMBIGUATING THE SCHEMA

Two of the ways our schema is ambiguous require little additional comment at this point. It is clear enough what it means to say that questions about metaphysics need at least initially to be distinguished from questions about explanatory success, though of course the relationship between certain types of explanation and metaphysics will need to be taken up at a later point. Likewise, it is clear enough what it means to distinguish claims made in the assertoric voice, as second-order empirical hypotheses about how the mind can be united with the natural sciences, from those made in the normative voice and intended to serve as a kind of constraint upon psychology or philosophy of mind. Of course, positive and normative claims must be evaluated in very different ways, and so it behooves us always to be careful in identifying which sort of claim we are dealing with.

By contrast, it is worth saying a bit at the outset about different views of what might be understood by 'the world of nature as understood by the natural sciences'. Some would-be 'naturalizers' of the mind are *reductionists*. Others are concerned with *lawlike* relations between mind and body, or among mental states. And still others wish to understand the mind in *biological* terms, employing resources from evolutionary theory or sociobiology. And these three approaches really reflect three different views of scientific explanation, which may be associated in turn with three important figures in the history of science: Galileo, Newton and Darwin. The following sections will thus explore, in turn, philosophical naturalisms built upon the adoption of one or another of these styles of explanation.

4. *Reductive Philosophical Naturalism*

For Early Modern mechanists like Hobbes and Descartes, the predominant model of explanation was styled upon mathematical demonstration: one understands how a complex object like a clock works by 'resolving' it into its component parts, postulating 'axioms' for how they move and interact, and then deriving or 'composing' the behavior of the whole from that of the parts. This 'method of resolution and composition' favored by Galileo and Descartes was an early form of reductive explanation. And this reductionist understanding of scientific explanation was directly tied to metaphysics. Descartes, for example, argued that animals are nothing but machines because (he claimed) everything they do can be explained by reducing

their bodies to mechanical components; but certain human faculties (reasoning and language) cannot be reproduced mechanically, and hence cannot be reduced to mechanical principles, from which he reasoned that what goes on in the body does not provide metaphysically sufficient conditions for thought. (In Cartesian terminology, a ‘conceptual distinction’ implies a ‘real distinction’.)

In the twentieth century, Logical Positivists and Empiricists were likewise inspired by the model of mathematical demonstration, and philosophers of science from Rudolf Carnap (*Aufbau*) to Ernest Nagel (*Structure of Science*) viewed scientific explanations within a domain as syllogisms and inter-theoretic explanations of the principles of the special sciences as reductions, in the form of axiomatic derivations of the special sciences from basic physics. It was within this philosophical climate that reductive physicalism became a popular view in philosophy of mind in the 1950s and 60s.

4.1. ANALYTIC BEHAVIORISM AND TYPE IDENTITY

Much of the reductionist theorizing of this period, however, was *not* styled upon the axiomatic model. The predominant notions of ‘reduction’ at that time were, instead, those of *analytic behaviorism and type-type identity*.

Analytic behaviorism is the thesis that claims made in the mentalistic vocabulary are *equivalent in meaning* to claims made in a vocabulary of stimuli and behaviors (Carnap, ‘Psychology in Physical Language’). The popularity of analytic behaviorism was comparatively brief. In part, it died out due to its failure to produce plausible non-mentalistic glosses on *any* part of the mentalistic vocabulary. And in part, it shared the fate of Skinnerian behaviorism, which was eclipsed by functionalism in philosophy of psychology (Putnam, ‘Brains and Behavior’) and by cognitivism, in the wake of the Chomskian revolution in linguistics, in psychology (Chomsky, *Aspects of the Theory of Syntax; Cartesian Linguistics*).

Type-identity theory is the thesis that that each legitimate mentalistic *kind* (e.g., pain) is *identical to* some (possibly complex) physical kind (e.g., C-fiber firings) (Place; Smart). This view was quite influential in the 1960s and 70s, but is now largely rejected by philosophers of mind. The main reason for its rejection was the growing popularity of functionalist accounts of mental state types in the 1970s and 80s. According to such accounts, mentalistic kinds, like ‘pain’ or ‘belief’, are typified *functionally*, much like biological kinds such ‘heart’ and kinds of computational circuits like ‘AND-gate’. Functional kinds are distinguished by *what* they do, rather than by *how* they do it, or by their structure. But functional kinds can be realized in multiple ways: human hearts and earthworm hearts bear few structural similarities, and circuits and programs with the same functional form can be implemented in computers that have little in common physically. So if mental kinds are functional kinds, they are *multiply realizable* – humans, earthworms, Martians, and androids might

have the same functional states, but these would be implemented differently in each case. The relation between mental and physical kinds is thus not one-to-one, as held by type-identity reductionists of the 60s, but one-to-many.

4.2. BROAD REDUCTIONISM

But the reductionist lineage descending from mechanism also gave rise to other views about the mind. The core of the mechanistic worldview was that (a) complex phenomena are completely a product of the interactions of their parts, and (b) a sufficient understanding of these interactions should allow us to derive the types of behavior we observe in the complex system. This 'broad reductionism' (Horst, *Beyond Reduction*) seems to be at work in a number of philosophers of mind, and consists in the view that we can give *part-whole* explanations of mental phenomena *without remainder* by appealing to neural, biological, or physical phenomena. Broad reductions are thus a type of *explanation*. The hallmark of such broad reductionism is the view that we understand a mental phenomenon only when we can fully explain it in non-mental terms, particularly terms appealing to the 'parts' out of which the mind might be 'built'.

Fred Dretske ('If You Can't Make One'), for example, seems to take such a view in advocating an 'engineer's' approach to cognition in his article, 'If you can't make one, you don't know how it works'. And Fodor also seems at times to embrace a reductionistic approach:

It's hard to see . . . how one can be a Realist about intentionality without also being, to some extent or other, a Reductionist. If the semantic and the intentional are real properties of things, it must be in virtue of their identity with (or maybe their supervenience on?) properties that are themselves *neither* intentional *nor* semantic. If aboutness is real, it must be really something else. (*Psychosemantics* 98)

Fodor is sometimes classified as an anti-reductionist on the grounds that he embraces functionalism. And this classification is correct, if by 'reductionism' one means *type identity theory*. But Fodor's functionalism really involves a commitment to two types of explanation that seem to count as broadly reductive. First, the nature of mental categories is supposed to be exhaustively analyzed in terms of their functional properties. Second, while functional *kinds* cannot be reduced to physical kinds, the fact that a given system counts as an instance of a particular functional kind is supposed to be fully explainable in terms of the interactions of the parts of the system, thus echoing Dretske's decompositional 'engineering' approach.

4.3. OTHER NOTIONS OF 'REDUCTION'

The word 'reduction' has also been used in additional ways within philosophy of mind. John Bickle (*Psychoneural Reduction; Philosophy and Neuroscience*)

has used the label ‘reduction’ for his more radical approach to philosophy of neuroscience. Bickle holds that philosophers of mind have erred in emphasizing explanations emerging from psychology and cognitive neuroscience, which attempt to treat the mind at the level of cognitive states and processes. His alternative is to look directly to neuroscience, and particular to processes taking place at a cellular and sub-cellular level, for a better understanding of the mind. He describes this as the ‘ruthless reduction’ of the mind to the phenomena studied by neuroscience. The result of such ruthless reduction, however, is not so much a vindication of the vocabulary and insights of common sense and cognitive psychology as a wholesale revision of them. (That is, ruthless reductionism has at least as much in common with eliminativism as with more familiar forms of reductionism in philosophy of mind.) Moreover, while Bickle originally saw this project as providing arguments for physicalism (*Psychoneural Reduction*), he has more recently taken it to imply a need to abandon traditional philosophical questions in favor of letting neuroscience be our guide to an understanding of the mind (*Philosophy and Neuroscience*). His project has thus become a form of *empirical* rather than *philosophical* naturalism.

4.4. REDUCTIONISM AS A PHILOSOPHICAL NATURALISM

Historically, reductionist approaches have enjoyed a kind of pride of place within philosophical discussions of naturalism. This is, in part, due to the fact that reductionism’s guiding metaphors – mechanism, decomposition, and axiomatic reconstruction – have enjoyed a long intellectual history. But both type identity theory and broadly reductionist approaches also have a unique relationship to important problems in the metaphysics of mind. Reductions of either sort are types of *explanation*. But they are types of explanation that have robust metaphysical implications. To reduce A to B is to explain why, if B is the case, A *must* also be the case. If we can reduce A to B, then we can *derive* A from B. And thus a reduction of A to B also provides proof that $B \rightarrow A$ is metaphysically necessary. A demonstration of type identity does even more than this: it shows the necessity of the biconditional ‘B iff A’. Thus a successful reduction of mental phenomena to neural or physical phenomena would constitute a strong argument for physicalism. This can be seen as a principle connecting explanation to metaphysics:

Positive Explanation-to-Metaphysics Connection Principle (Positive EMC): If A is reducible to B, then $B \rightarrow A$ is metaphysically necessary, and A is metaphysically supervenient upon B.

4.5. THE APPEAL OF REDUCTIONISM

Broad reductionism is still a mainstream view in philosophy of mind. It is viewed as an attractive position for a number of reasons. Two of these

have already been mentioned: First, there is a long-standing history in philosophy of science in which reductive explanation, styled upon mathematical demonstration, was viewed as *the* paradigm of scientific explanation. Second, reductive explanations allow for a direct argument to metaphysical conclusions. And indeed, it is not clear that anything *short of* reductive explanations can provide a demonstration of the truth of physicalism.

But there are also two additional factors that have contributed to the plausibility of reductionism. One of these is the assumption that, *outside* the sciences of the mind, the special sciences (such as chemistry and biology) have generally proven to yield to reductions to fundamental physics. This does not *entail* that mental phenomena will similarly prove to be reducible to something non-mental. But it does suggest a general picture of nature in which things other than basic physical phenomena are reducible to physics. And if this is so, then one of two things must be the case: either mental phenomena are also reducible, or else they are radically discontinuous with (the rest of) the natural world. This has given rise to *normative* arguments for reductive naturalism. Some of these are *inductive* in character: because the general pattern we find is that mature special sciences turn out to be reducible to physics, we should expect that the same will prove to be the case with mental phenomena, as they are better understood by the sciences. Others are more directly normative: unless mental phenomena can be reduced to something non-mental, we have some reason to doubt the legitimacy of mentalistic claims altogether. That is, there is a kind of forced choice between reduction and elimination of mental phenomena (cf. Stich, 'What is a Theory of Mental Representation?').

An additional motivation for reductive naturalism has come from advances in the sciences of the mind, and particularly from the discovery of many robust relationships between mental phenomena and things going on in the brain. To the extent that these may be viewed as *reductive* explanations, there is reason to think that the cognitive sciences are even now providing direct evidence that mental phenomena are reducible to processes in the brain.

4.6. PROBLEMS FOR BROAD REDUCTIONISM

Despite all this, enthusiasm for reductive naturalism has begun to wane over the past decade. The most influential problem for reductive naturalism is the apparent existence of an 'explanatory gap' between mind and brain (Levine; Jackson, 'Epiphenomenal Qualia', T. Nagel; Chalmers). Proponents of the explanatory gap differ on the exact list of mental states that are supposed to be irreducible. Some concentrate on *qualia* and consciousness. For example, Chalmers argues that many psychological state types (such as belief and desire) are functionally defined, and hence are candidates for (broad) reduction, but that *qualia* and consciousness are defined neither

structurally nor functionally, and hence are not apt candidates for reduction. Others, like Searle, Horst (*Symbols, Computation and Intentionality*), Siewert, Brandom (*Making it Explicit*), and Horgan and Tienson, include intentional states among those that are irreducible, either because they are intrinsically bound up with consciousness and intentionality (thus falling under arguments similar to that offered by Chalmers) or because they are intrinsically normative (thus falling under arguments similar to those of ethical anti-naturalists) (Putnam, 'Why Reason Can't be Naturalized'; Brandom, *Making it Explicit*; 'Modality, Normativity, and Intentionality').

Some advocates of the explanatory gap have gone further, arguing that the existence of such gaps provides reason to reject, not only *reductive* physicalism, but *any* form of physicalism, in favor of dualism. We noted earlier that reducibility entails metaphysical supervenience. Conversely, if a mental phenomenon is *irreducible*, there is something about it that is left unexplained by the neural or physical facts. Opponents of physicalism, particularly dualists, have argued that such a failure of reducibility entails a failure of metaphysical supervenience as well, and hence a need to hold that there are either non-physical *substances* or at least non-physical *properties*. Dualists have thus traded heavily upon a second, and more controversial, principle relating explanation to metaphysics:

Negative Explanation-to-Metaphysics Connection Principle (Negative EMC): If A is not reducible to B, then 'B and not-A' is metaphysically possible, and A does not supervene metaphysically upon B.

Negative EMC is rejected by non-reductive physicalists, who take the 'Mysterian' view that there can be metaphysical supervenience relations that whose necessity is epistemically opaque to us, due to some limitation upon our cognitive capacities (McGinn; Pinker). Some non-reductive physicalists are thus advocates of a *metaphysical* naturalism without a thoroughgoing *explanatory* naturalism. (Davidson, for example, holds that each mental event is token-identical with a physical event, but that there are not even nomic connections between physical and mental types. See below.) Others appeal to some other, non-reductive, form of explanation in order to be explanatory naturalists as well (see Sections 5–6).

Debates about the status and implications of the explanatory gap between mind and brain have occupied center stage in recent philosophy of mind, involving reductive and non-reductive physicalists, dualists, and eliminativists. Over the past decade, there has been a growing acceptance of the claim that the sciences of the mind are *not* thus far providing broad reductions of consciousness, intentionality or normativity, and that there indeed seem to be abiding explanatory gaps around these phenomena. In general, however, all parties have assumed that, *except for* certain mental phenomena, the objects of the special sciences are generally subject to inter-theoretic reduction. But Horst (*Beyond Reduction*) has pointed out that, during the very period that such debates were thriving, this assumption

was undergoing decisive critique within philosophy of science. In fact, inter-theoretic reductions are quite *rare*, and the absence of reductions is not generally taken to threaten either the metaphysical or the methodological credentials of sciences like chemistry and biology (see also Stich and Laurence; Baker). If this is the case, then it is, as it were, ‘gaps all the way down’, at least if any failure of reducibility counts as an explanatory gap.

This is a direct blow to reductive physicalism, and also to those forms of eliminativism that treat reducibility as a kind of litmus for ontological or scientific respectability. If inter-theoretic reductions are rare, this undercuts the inductive argument for psycho-physical reduction based upon the examples of other sciences. It also undercuts normative arguments for psycho-physical reduction: if inter-theoretic reductions are rare, then reducibility ought not to serve as a litmus for ontological or scientific respectability. But post-reductionist philosophy of science also presents problems for dualists and non-reductive physicalists. Dualists have traded heavily upon Negative EMC to turn the explanatory gap between mind and brain into a case for the view that mental phenomena do not supervene upon physical phenomena. But if biology and chemistry are likewise irreducible to physics, Negative EMC would entail that biological and chemical phenomena do not supervene upon physical phenomena either. This might be a hard pill for even dualists to swallow. At very least, it undercuts the dualist argument that the relation of the mental to the physical is metaphysically unique. But if they address this problem by disavowing Negative EMC, they also deprive themselves of the major argument for dualism. Non-reductive physicalism might appear to fare better, as non-reductive physicalists reject Negative EMC and are already prepared to acknowledge that supervenience relations may be abidingly inscrutable. But in fact, the *general* absence of reductions leaves the physicalist with no decisive reason to prefer a physicalist metaphysics to its alternatives.

4.7. NON-REDUCTIVE PHYSICALISM

It may seem peculiar to include a subsection on non-reductive physicalism within a larger section on reductive naturalism. The section organization of this paper, or at least of Sections 4 to 6, is structured according to types of scientific *explanation* – reductive, nomic, or evolutionary – favored by different types of philosophical naturalists. Non-reductive physicalists, as a group, do not share a single view of the type of scientific explanation through which the mental is to be accommodated within the natural sciences. Indeed, some deny that crucial mental properties like consciousness or intentionality *can* be satisfactorily explained. Non-reductive physicalism has been defined in large measure by contrast with reductive physicalism; and so there is some method to my apparent madness in turning to it here (see Melnyk).

There are at least three different reasons that philosophers have been drawn to non-reductive physicalism over the past several decades. One reason for its rise has been the influence of Donald Davidson's anomalous monism. Followers of Davidson hold that each mental state is token-identical with some physical state, while also holding that even a complete description of the physical world does not uniquely determine the content of the mental state in question, due to the indeterminacy of interpretation. This strikes me as a thoroughly non-reductionist account of the mental on two grounds. First, physical states underdetermine intentional states. Second, what makes a mental state mean anything is not limited to the physical states it is token-identical with, but requires the addition of an interpretation in accordance with a rationality-ascribing principle of charity. However, the appeal to non-physical principles of rational interpretation might call into question whether Davidson's anomalous monism deserves to be called naturalistic. It is naturalistic in its ontological inventory: every phenomenon (including each act of ascribing intentional content to others or oneself) is token-identical with some physical phenomenon. But Davidson's view is not purely naturalistic in its *principles*, as it invokes rational norms over and above the physical phenomena that are found in the sciences. To the extent that the one views the semantic and rational normativity in Davidson's account to be itself unnaturalized, one might view it as less than a thoroughgoing naturalization of the mind.

A second strand of non-reductive physicalism arose from discussions of functionalism and multiple realization. Unlike Davidson, functionalists generally hold that the realization relation between functional states and physical states does *not* require additional rational norms. Each token functional state is the state it is simply by dint of the fact that its physical realization base has the causal properties necessary to count as implementing said function. Indeed, some functionalists would go further than this, and hold that the relation between functional and physical states does not lie at the level of individual tokens, but of *kinds* of functional systems. For example, pain might be realized differently in humans, Martians and androids, but many functionalists assume that it is realized in the same way in all humans. When the fact that a system's meeting the criteria for a functional description F can be derived from its physical properties P, this amounts to a 'local' reduction – say, of pain-in-humans – of F to P, even though there is not a single reduction of pain to a common set of physical phenomena that encompasses all of the instances (cf. Kim, 'Multiple Realization'). Such a view strikes me as really being a form of broad reductionism. The type of 'reduction' it rejects is type-identity.

A third strand of non-reductive physicalism arises from an acceptance of the idea that there may indeed be explanatory gaps between mind and brain that are unsurmountable, combined with a commitment to physicalism as an ontological thesis. This seems to be the 'Mysterian' view of Colin McGinn. Mysterianism is the view that there are phenomena that the

human mind is unsuited to understanding. We might further divide Mysterianism into two varieties: a 'Strong Mysterianism', which holds that some phenomenon X cannot be understood, *even in its own right*, and a 'Weak Mysterianism', which holds merely that X cannot be completely *understood in terms of something else*. Mysterianism about the mind is a thesis independent of physicalism. Dualists, after all, are at least weakly Mysterian, in that they hold that the mental cannot be completely explained in non-mental terms. Non-reductive physicalists of the Mysterian sort believe that each mental phenomenon M is supervenient upon some physical phenomenon P, but that the $P \rightarrow M$ connection is not one that we can fully comprehend.

The Mysterian half of this view seems well-motivated by the difficulties in bridging the explanatory gap between mind and brain. But we would do well to ask what, given the lack of reductive explanation of the mental, is the reason for believing that mental states even *supervene* upon physical states or brain states. Historically, I suspect that the line of reasoning goes like this. Non-reductive physicalists of the Mysterian stripe generally started out as reductive physicalists, and then became convinced that reductions of consciousness and intentionality were not forthcoming. Non-reductive physicalism offered itself as an avenue of strategic retreat, giving up only what seemed unavailable in any case. And this alternative might seem quite reasonable, assuming that, *apart* from mental phenomena, everything else in the world is clearly supervenient upon basic physics because of the availability of inter-theoretic reductions of the other special sciences. However, given the turn against inter-theoretic reduction in philosophy of science mentioned in the previous section, it is not clear that this move is available. If the unavailability of such reductions means it is 'explanatory gaps all the way down', then one must be, in some sense, a 'Mysterian all the way down' as well. If there is a general lack of proof that even *natural* phenomena (of the sort discussed in chemistry or biology) are metaphysically supervenient upon physical phenomena, this considerably weakens the inductive argument that mental phenomena are likely to be metaphysically supervenient upon physical states. The conjunction of Mysterianism with physicalism is *consistent*, but it stands in need of some other type of motivation for its physicalist commitment.

As a result, some physicalists nowadays argue for physicalism on other, non-reductive grounds. In particular, it is claimed that the universe is causally closed under physics, and hence that there are no phenomena with causal powers that do not supervene upon physical phenomena (e.g., Papineau, *Philosophical Naturalism*). However, one would do well to ask whether we in fact know that the universe is causally closed under physics. In *Thinking about Consciousness*, Papineau indicates that he had originally thought that the thesis of causal closure under physical laws was

not a problematic issue . . . The one assumption that I did not expect to be controversial was the completeness of physics. To my surprise, I discovered that

a number of my philosophical colleagues did not agree. They didn't see why some physical occurrences, in our brains perhaps, shouldn't have irreducibly conscious causes. My first reaction to this suggestion was that it betrayed an insufficient understanding of modern physics. Surely, I felt, the completeness premise is simply part of standard physical theory. However, when my objectors pressed me, not unreasonably, to show them where the completeness of physics is written down in the physics textbooks, I found myself in some embarrassment. Once I was forced to defend it, I realized that the completeness of physics is by no means self-evident. Indeed, further research has led me to realize that, far from being self-evident, it is an issue on which the post-Galilean scientific tradition has changed its mind several times. (45)

We would do well to take Papineau's lesson to heart. Indeed, while Papineau's admissions on this subject reflect admirable intellectual honesty, his own attempts to address the issue in the appendix to *Thinking about Consciousness* are curiously unsatisfying. There, he observes that, by about 1900, there were two areas of scientific inquiry in which there was serious question about whether it was necessary to include nonphysical causal principles: living systems and the conscious mind. He then provides a brief overview of developments in the life sciences over the course of the twentieth century that unlocked a number of mechanisms underlying processes like metabolism, and thereby stripped vitalism of much of its previous allure. But – and this is a curious fact given that the book in question is a book about consciousness – he never addresses, in similar fashion, how far neuroscience might go in addressing cognate concerns about the mind. Instead, he seems to argue by *analogy*: that because developments in the life sciences have progressively provided alternatives to the vitalist impulse, it is reasonable to expect that future work in neuroscience or other disciplines will do the same with respect to the antiphysicalist impulse. This argument is quite unsatisfying. On the one hand, it does nothing to engage dualists, who hold that mind-brain connections are *disanalogous* with chemistry-physics connections in just this regard. On the other hand, one might question just how far inter-theoretic connections between physics and the life sciences really go towards demonstrating causal closure.

In summary, non-reductive physicalism of the Mysterian sort is a consistent position, but its commitment to physicalism strikes me as ultimately an expression of philosophical taste. Non-reductive physicalism of the sort embraced by many functionalists is actually a species of broad reductionism. And Davidsonian non-reductive physicalism relies on non-physical principles of rational interpretation, which arguably call into question its status as naturalistic.

5. *Nomic Philosophical Naturalism*

In the 18th century, followers of Isaac Newton tended to reject the reductionistic model of scientific explanation favored by the mechanists.

In an oft-cited passage in the General Scholia to the 2nd edition of the *Principia*, Newton wrote

Hitherto we have explained the phenomena of the heavens and of our sea by the power of gravity, but have not yet assigned the cause of this power. . . . But hitherto I have not been able to discover the cause of those properties of gravity from phenomena, and I frame no hypotheses; for whatever is not deduced from the phenomena is to be called an hypothesis; and hypotheses, whether metaphysical or physical, whether of occult qualities or mechanical, have no place in experimental philosophy. In this philosophy particular propositions are inferred from the phenomena, and afterwards rendered general by induction. Thus it was that the impenetrability, the mobility, and the impulsive force of bodies, and the laws of motion and of gravitation, were discovered. And to us it is enough that gravity does really exist, and act according to the laws which we have explained, and abundantly serves to account for all the motions of the celestial bodies, and of our sea. (546–7)

Exactly what Newton himself understood by ‘I feign no hypotheses’ (*‘hypotheses non fingo’*) is a matter of lively scholarly debate. However, Newtonians like Hume took this to license a rejection of the search for hypothetical unseen mechanisms in favor of a search for mathematical laws that describe the observable phenomena. Science was not to be in the business of postulating mechanisms so much as finding laws that would allow for prediction and control.

There are four junctures at which one might look for laws involving mental states:

- 1) lawful relations between two or more mental states;
- 2) lawful relations between stimuli and mental states;
- 3) lawful relations between mental states and behaviors;
- 4) lawful relations between brain states and mental states.

5.1. LAWS LINKING MENTAL STATES TO ONE ANOTHER

The attempts of 18th- and 19th-century ‘Newtonians’ to extend Newtonian methodology to psychology, from John Locke to James Mill, tended to eschew the search for mind–body connections (like Hobbes’s sketchy identification of sensations and appetites with motions in the body) in favor of a ‘mental chemistry’ that would uncover laws (generally understood to be laws of association) linking one mental state to another.

Recent philosophers have sought to vindicate intentional psychology on the grounds that it (either in its ‘folk’ form or regimented into scientific cognitive psychology) trades in psychological laws, relating mental states, not to their neural correlates, but to other mental states. Indeed, functionalism seeks either to define or to explicate mental state types in terms of such relations, as well as relations to stimuli and behaviors. Much of the debate between eliminativists (Stich, *From Folk Psychology to Cognitive*

Science; Churchland, 'Eliminative Materialism and the Propositional Attitudes') and functionalists (e.g., Fodor, *Representations*) in the 1980s centered around the question of whether the generalizations of folk psychology are (or can be made into) *bona fide* laws. Here, the presence of such laws served a kind of normative role for the scientific status of the human sciences, harkening back to an older debate between Logical Positivists and the *Verstehen* tradition concerning whether the human sciences deal in laws, like the natural sciences, or instead deal (merely) in interpretations. Many projects in artificial intelligence and cognitive psychology inspired by the computer metaphor were attempts to make explicit the sort of processes that could underwrite such lawlike relations between intentional states, and philosophers used the successes of the models they produced in an attempt to vindicate the status of intentional psychology. Such projects were generally viewed as naturalistic on the assumption that the mind actually makes transitions between intentional states by implementing the types of computational mechanisms posited by cognitive psychologists and AI programmers.

Other philosophers have also developed versions of naturalism which treat participation in laws as a necessary and perhaps sufficient condition for mental states to count as naturalistic. Michael Tye, for example, proposes that naturalism in philosophy of mind be understood as the view that: '[m]ental states participate in causal interactions which fall under scientific laws, and are either ultimately constituted by or ultimately realized by microphysical phenomena' (436). Tye takes it as unproblematic that mental states fall under laws in this way, and that this fact implies that we do not need to naturalize such states, because they are *already* 'natural' by dint of participating in lawful relationships. Given that he appeals to psychology rather than psychophysics or neuroscience, one may assume that he is thinking primarily of inter-mental laws:

Psychology is a science no different in its procedures and laws from other sciences. So, of course, the mental is a part of nature in the ways I have described. To suppose otherwise is to suppose that there is something peculiar about the mental which prevents it from having the features adumbrated above. And there just is no good reason for any such supposition. So, as I commented in the introduction to this paper, naturalism with respect to the mental, once properly explicated, is really beyond question. (437)

Tye's account also requires that mental states be 'either ultimately constituted by or ultimately realized by microphysical phenomena'. However, the notions of 'constitution' and 'realization' are not fully explicated, beyond the fact that they are part-whole relationships. Instead, Tye introduces them by comparison with the relations between other special sciences and physics:

Just how is naturalism with respect to the mental to be explicated? The answer, I suggest, is that, for the naturalist, mental states should not only participate in

causal interactions describable in scientific laws but also bear the same general ontic relationship to lower level physical items as do the physical entities quantified over and referred to in higher level physical laws generally (for example, those in biology and geology). (434)

Given Tye's critical dismissal of Fodor's view that a naturalization of a mental phenomenon must provide a sufficient condition for it in non-mental terms (429), we must presume that his notions of 'constitution' and 'realization' need not supply metaphysically sufficient conditions for mental states. I am thus inclined to regard at least this part of Tye's naturalism as a form of *empirical* naturalism – that is, a turn towards the sciences that replaces traditional metaphysical questions rather than providing answers to them. Alternatively, his may be a mixed account of the strictly philosophical sort, one which accords both participation in laws and physical composition a role in determining whether a philosophy of mind is naturalistic, while rejecting a need for broad reduction.

5.2. LAWS RELATING STIMULI TO MENTAL STATES

In the 19th century, such a law-centered vision of science was extended to the connections between percepts and the stimuli that cause them by psychophysicists like Weber and Fechner, and also by crossover work by physicists like Helmholtz. Such psychophysical laws play an important role in Fodor's attempt to 'vindicate' psychology (*RePresentations*).

There has been a limited amount of discussion, by both scientists and philosophers, about whether there really are psychophysical laws (Savage, *Measurement of Sensation*; Laming), and if so, whether they relate stimuli to phenomenological states. Many interpreters of psychophysics have viewed things like the Weber–Fechner laws as reporting only discriminative capacities, and not relations between intensities of stimuli and intensities of subjective percepts (e.g., Ratliff; Ratliff and Sirovich). This may be a tenable view of the parts of psychophysics dealing only with intensities and with discrimination of other stimulus properties like hue. However, it seems far less plausible with respect to many other psychophysical data, such as visual illusions where the *profile* of intensities across a percept is an important datum (Todorovic) or where the percept involves the perception of a figure, and hence intentionality, particularly in cases like subjective contour figures (Horst, 'Phenomenology and Psychophysics').

5.3. LAWS RELATING MENTAL STATES TO BRAIN STATES

Several other contemporary philosophers speak of the relationships between mental states and their corresponding brain states as 'psychophysical laws' as well, albeit in a different sense from the 'psychophysical laws' discovered by Weber and Fechner (e.g., Davidson; Chalmers). Laws of the sort discussed by Weber and Fechner relate stimuli to mental states. What

Davidson and Chalmers have in mind, by contrast, are relations between mental states and their 'neural correlates', the lawfulness of the relationship consisting, at very least, in this correlation. (Fechner called this 'inner psychophysics' in contrast with the 'outer psychophysics' typified by the laws he explored.)

Davidson argues, with respect to beliefs and desires, that *are no* such psychophysical laws relating them to brain states, or even to states of the entire physical universe. More exactly, he argues that a specification of any physical state (including specifications in neural or behavioral terms) does not uniquely determine an intentional state, on the grounds that the content of intentional states is not intrinsic, but relative to an interpretation, and there are always multiple consistent interpretations available. Each mental state is token-identical with *some* physical state, but there is no lawlike function from a physical state to a unique mental state. Davidson's position is known as 'anomalous monism' because it embraces token identity while denying psycho-physical laws. It became a powerful force in motivating the rise of non-reductive physicalism in the 1980s. However, it goes beyond a denial of reduction in denying that there are even nomic psycho-physical relations. It is important to note that the type of 'mental' states Davidson has in mind are beliefs and desires, and not the kinds of experiential percepts that play a role both in Fechner's psychophysics and in discussions of consciousness, like that taken up by Chalmers. One could adopt an anomalous monism with respect to intentional states while embracing a reductive physicalism with respect to sensations and other qualia.

Chalmers, in his investigation of consciousness, posits that there *are* psycho-physical laws relating brain states, or more generally physical states, to conscious mental states. But he argues that the conscious mental states are not *reducible* to the physical states to which they are lawfully related. In *The Conscious Mind* (1996), he held that this stands in sharp contrast with all other types of phenomena, including mental states other than qualia, which he claimed to be functionally typed, and hence subject to reductive explanation. Subsequently, Chalmers has expressed openness to the view that intentional states involve consciousness, and hence fail to be reducible on the same grounds. Chalmers embraces property dualism, and one might think that this would lead him to style himself an anti-naturalist. However, he in fact drapes his position in the mantle of naturalism, treating non-reductive lawful relations between mind and matter as sufficient for naturalism. He writes, for example, 'The third constraint is that I take consciousness to be a natural phenomenon, falling under the sway of natural laws' (xiii). It is not clear whether the clause 'falling under the sway of natural laws' is supposed to function as an explication of what it is to be a 'natural phenomenon' or whether it is an additional claim. But his is clearly intended to be a nomic philosophical naturalism.

5.4. THE PROSPECTS OF A NOMIC NATURALISM

If ‘naturalizing’ the mind consists in finding laws relating (a) pairs of mental states to one another, and/or (b) mental states to stimuli, and/or (c) mental states to behaviors, and/or (d) mental states to brain states, the naturalist has an agenda that is considerably less ambitious than broad reduction. *Reductive* connections, modeled on mathematical demonstration and construction, carry the force of metaphysical necessity. By contrast, *laws*, even physical laws, are generally held to be metaphysically *contingent*. They consist first and foremost in robust empirical generalizations about things that co-occur, and often involve the postulation of causal connections between the events related by the law. (And when this is so, it is not a reduction, as causation is a relation between two distinct events rather than a relation between the parts and the whole.) Finding an empirical generalization that relates A and B does not preclude a reduction of A to B, but it does not entail it either; and indeed if such a reduction is found, we might well cease to speak of the relation as a law.

A merely nomic (that is, lawful) connection between mind and body, however, is compatible with a variety of metaphysical interpretations. It is compatible with materialism. But it is also compatible with property and substance dualism, and for that matter with various forms of idealism, pragmatism, neutral monism, and social constructionism as well. Fechner was in fact hostile to materialism, and viewed his ‘outer’ psychophysical relations as relating physical stimuli and phenomenological percepts. Correlations between mental states and their neural correlates *might* be a result of identity or metaphysical supervenience, but they might likewise be a result of a contingent set of psycho-physical laws, as Chalmers suggests. Lawful causal relations between intentional states might be emergent from the dynamics of the underlying physical and neural systems, but the existence of such laws is also compatible with the possibility of a special and independent form of mental-mental causation.⁵

This would seem to violate the Caveat offered earlier, that disqualifies views that countenance things like Cartesian souls, or that do not treat the lowest-level categories of the natural sciences as fundamental, from being labeled ‘naturalistic’. One can, of course, choose to use the word ‘naturalism’ in a weaker sense, as for example Chalmers does. In part, this is merely a dispute over words; but this usage seems to go against the spirit and motivations associated with the use of the word, both historically and on the contemporary scene. And so I shall take the view that *a merely nomic form of naturalism is really not a species of philosophical naturalism at all*, especially as it could have been endorsed by someone like Descartes, so often identified as a principal and even paradigmatic opponent of naturalism, who thought there were nomic *causal* relations between mind and body.

6. Evolutionary Philosophical Naturalism

Still other proponents of views styled ‘naturalistic’ are interested in accommodating the mental under the aegis of evolutionary biology. In its mildest form, Darwinian naturalism treats specific types of mental states – pains, desires, beliefs – as phenotypic features of an organism that are to be explained through mechanisms of variation and selection at work in the ancestral history of the species. There are four elements to such an evolutionary story, consisting in accounts of: (1) how the trait initially comes on the scene through some process of spontaneous variation; (2) how it is heritable from one generation to another; (3) how it is expressed in an individual organism through development; and (4) how mechanisms of selection account for its proliferation as an adaptation. Evolutionary psychology generally concentrates upon the final element, telling stories about the hypothesized adaptive value of various mental traits (e.g., essays in Barkow, Cosmides, and Tooby).

Some forms of Darwinian naturalism go further than this. Millikan (*Language, Thought and Other Biological Categories; On Clear and Confused Ideas*), for example, attempts to account for the *nature* of mental traits through biological explanation. The nature of a trait is its proper function. Likewise Dretske (*Naturalizing the Mind*) differentiates what a mechanism in an organism *actually does* from its *function*, understood as what it was *selected* to do. These projects, while they draw upon ideas from evolutionary biology, are primarily philosophical projects. Other philosophers have drawn more directly upon empirical research in evolution and the life sciences, and will be briefly canvassed in Section 7, on empirical naturalism.

Evolutionary explanation is sometimes viewed as closing an important gap between physics and psychology, and thus as providing a necessary supplement to reductive explanation. However, this is misleading. Stories about the adaptive value of a phenotypic trait bring that trait fully within the broader scope of the natural world only when supplemented with the rest of the evolutionary story about the appearance, inheritance, and expression of the trait. Consider two extreme examples. An organism that was supplemented with a Cartesian rational soul would likely enjoy competitive advantages over organisms lacking such a soul, because it would confer language and rationality upon its bearer. You could tell a good story about how Cartesian souls would be *adaptive*. However, having an immaterial rational soul is not the sort of thing that could be transmitted genetically to one’s offspring, and hence not a trait on which gene selection could operate. Likewise, it is not the sort of thing that could be the result of the expression of genes. Or consider a second example: an organism that was powered by a perpetual motion machine would have an enviable degree of differential fitness in that it would not need to eat to live, and hence would be immune to famine and could devote more of its energies to producing offspring. However, we have good reason to

suppose that the physical world could not endow an organism with a perpetual motion machine, and hence could not supply the preconditions for forces of selection to operate (Horst, 'Evolutionary Explanation').

The moral of these (admittedly extreme) stories is that evolutionary explanations are suspect in precisely the cases where there is reason to wonder whether merely physical mechanisms *could* indeed produce the phenotypic trait in question. Settling the question of whether physical mechanisms can do so in a given case is precisely what is at stake in reductive explanations. To the extent that one has reason to doubt that a mental trait is indeed subject to reduction, one thereby has reason to doubt that it is something that could arise through mutation, be expressed through development, or upon which mechanisms of selection could operate. And so, for example, arguments to the effect that mental phenomena like consciousness and meaning *cannot* be accounted for by the physical phenomena going on in the brain are by extension arguments against evolutionary accounts of the mind as well. Concentrating on the *selectional* component of evolutionary explanation creates the illusion of bypassing the problems of alleged explanatory and metaphysical gaps; but an illusion it is. We cannot tell a story about the inheritance or selection of a trait unless it is something that could be the result of the expression of genes in development, and passed on through physical mechanisms of inheritance, and these are precisely what such anti-naturalistic arguments call into question.

The issue here is different from that encountered by nomic forms of naturalism. There, the issue was that laws, even universal laws, relating mental and physical states are compatible with paradigmatically *non-naturalistic* interpretations. In the case of evolutionary accounts, the issue is that the selectional stories that are often told work as *naturalizations* only on the assumption that the other parts of a complete evolutionary account could, in principle, be filled in. The non-naturalist can allow that such accounts *might* explain mental phenomena *if* the rest of the evolutionary story can be assumed, but doubt that it can safely be assumed. And in particular, insofar as there are problems for attempts to reduce mental phenomena to physical phenomena, there is likewise reason to doubt that the rest of the evolutionary story *can* safely be assumed.

7. Empirical Naturalisms

The 'naturalistic' projects thus far surveyed – varieties of what I have called 'philosophical naturalism' – have been driven in large measure by traditional issues in philosophy of mind: the nature of mental states, physicalism vs. dualism, the relation between mind and body, and the search for sufficient non-mental conditions for mental states. Philosophical naturalists often make appeals to science, but their projects are distinctively philosophical rather than scientific. Indeed, many philosophical naturalists make little use of actual scientific studies at all.

But the label ‘naturalism’ has also been used by philosophers who see the relation between philosophy and science in a very different way. For them, the concrete results of the sciences – what they have to say about matter or mind – represent the primary discourse, and philosophy plays only a supporting role, by commenting upon the sciences, or by providing initial pre-scientific explorations of problems until such time as they can be addressed more adequately by a mature science. These ‘empirical naturalists’ reject the views that philosophy has a distinctive set of methods, particularly *a priori* methods, and that it can play a special authoritative role legislating what is ‘good science’. An empirically naturalistic philosophy of mind is thus some combination of philosophical commentary upon the sciences of the mind and philosophical exploration of mental phenomena in the hope that they will eventually become the domain of mature sciences.

The ‘naturalism’ endorsed here is quite similar to what is meant by ‘naturalism’ in epistemology and philosophy of science. Quine’s ‘Epistemology Naturalized’ (1969) is often regarded as the seminal statement of this type of naturalism. Quine writes that ‘Naturalism looks only to natural science, however fallible, for an account of what there is and of what what there is does’ (*Pursuit of Truth* 9). Alex Rosenberg, in a survey article on naturalism, expands upon Quine’s theme:

[W]e may characterize naturalism in philosophy as follows:

1. The repudiation of ‘first philosophy’. Epistemology is not to be treated as a propaedeutic to the acquisition of further knowledge.
2. Scientism. The sciences—from physics to psychology and even occasionally sociology, their methods and findings, are to be the guide to epistemology and metaphysics. But the more well-established the finding and method the greater the reliance philosophy may place upon it. And physics embodies the most well-established methods and findings. (4)

Rosenberg’s article is aimed primarily at characterizing naturalistic philosophy of *science*, but applies as well to approaches to the mind that are empirically naturalistic. (To this list, Rosenberg adds a third characteristic of ‘Darwinism’. However, it seems clear that there are many self-styled naturalists whose approach is little influenced by evolutionary theory.)

A number of philosophers of mind echo Quine’s view at a programmatic level in their glosses on ‘naturalism’. Peter Godfrey-Smith writes: ‘Contemporary naturalism stresses the continuity of philosophy with science’ (8). Michael Devitt also endorses this approach, writing that ‘there is only way of knowing: the empirical way that is the basis of science (whatever that way might be)’ (45). And Lawrence Shapiro takes a similar view, which he calls ‘methodological naturalism’ in contrast with what he somewhat disparagingly calls ‘LEGO naturalism’ (basically, a reductionist metaphysical naturalism), recommending that we ‘call “natural” those things amenable to scientific scrutiny’ (318). And Kim Sterelny

writes, 'naturalists . . . think philosophy is continuous with the natural sciences. On this view, philosophical theories are conjectures whose fate is ultimately determined by scientific investigation' (*Representational Theory of Mind* qtd. in Jackson, *From Metaphysics to Ethics* 155).

This view that philosophy of mind is really a kind of proto-science is fleshed out in concrete projects by philosophers exploring the resources afforded by particular scientific disciplines. Natural candidates for exploration include cognitive and developmental psychology (Jackendoff; Lakoff; Posner; Goldman; Boterill and Carruthers; Lakoff and Johnson; Chomsky, *Language and Mind*) and neuroscience (P. S. Churchland; Bickle, *Psychoneural Reduction; Philosophy and Neuroscience*; Hardcastle). But they also include a number of disciplines within biology: cognitive ethology (Allen and Bekoff; Carruthers), evolutionary theory (Carruthers; Sterelny, *Thought in a Hostile World*; Barkow, Cosimedes, and Toobey), and environmental complexity (Godfrey-Smith). One might also include in this group explorations of more formally-oriented disciplines like dynamic systems theory (Maturana and Varela; Port and Van Gelder; Varela, Thompson, and Rosch) computer science and artificial intelligence (e.g., essays in Pollock and Cummins; Haugeland), though some empirical naturalists might question the credentials of computer science and AI as *empirical* as opposed to *formal* disciplines.

8. Do Empirical and Philosophical Naturalism Make for Good Bedfellows?

What is the relation between empirical naturalism and philosophical naturalism? There is not one single answer, and different philosophers have taken strikingly different positions on this question. Some have viewed them as natural bedfellows: whether because they view empirical naturalism as providing support for some type of philosophical naturalism, or because they believe that philosophical naturalism suggests that real progress in understanding the mind will come through the sciences rather than philosophy. Others, following Quine, have taken the view that adopting the attitude of empirical naturalism should lead to a rejection of most or all traditional projects in metaphysics and epistemology: science is the arbiter of what exists, and knowledge is itself a phenomenon to be understood through the cognitive sciences. And still others have argued that empirically naturalistic accounts of the mind either leave the issues between philosophical naturalists and their critics untouched, or else are harmful to the agenda of at least some forms of philosophical naturalism.

Does philosophical naturalism have any implications for whether one should pursue the investigation of the mind as an empirical naturalist? Some answer *yes* to this question. But there are also a number of philosophical naturalists who are engaged primarily or even exclusively in traditional projects in metaphysics or epistemology, and whose work has little real engagement with the sciences of the mind. Indeed, one can endorse

naturalistic views like reductive physicalism on entirely philosophical grounds. Jaegwon Kim, for example, argues for physicalism on the basis of arguments to the effect that, if mental states do not supervene upon physical states, one is faced with a problem of 'double causation' – that mental states and actions have both completely adequate physical causes *and* an additional set of mental causes (*Supervenience and Mind*). This is a purely philosophical argument, based on *a priori* reasoning of the sort rejected by empirical naturalists. So a commitment to at least some types of philosophical naturalism is consistent with holding that there is also a set of distinctively philosophical questions and methods that are relevant to the project of understanding the relation between mind and body.

Does empirical naturalism have any consequences for philosophical naturalisms of various sorts? What is at the core of empirical naturalism is a kind of methodological maxim: to understand the mind (and epistemology) through the resources of the sciences. The result of this approach will inevitably be to restrict our theories of mind to what is found in the sciences themselves. This may indeed screen out some possible lines of philosophical inquiry, such as speculations about immaterial souls that cannot be studied in the laboratory through controlled interventions. And if the sciences of the mind were to produce plausible reductions of mental phenomena, this would indeed lend credence to reductive physicalism. But if such reductions are not forthcoming, the morals of empirical investigation are less clear. For example, if there are lawlike relations between mental states, or between mental states and neural states, this is compatible with non-physicalist views in metaphysics. The empirical naturalist could well hold that metaphysical questions about the mind are legitimate questions, even if they are scientifically intractable, and indeed even if they cannot be definitively settled. On the other hand, one could strengthen the canons of empirical naturalism further, to hold that questions that cannot be settled by science are somehow *illegitimate questions*. This, however, would categorize as nonsense, not only metaphysical claims about the mind, but also a whole range of other non-scientific claims that we intuitively feel are both sensible and have truth values. This seems a very high price to pay.

Empirical naturalism can, however, have some *negative* implications for certain forms of philosophical naturalism. In particular, *normative* forms of philosophical naturalism, such as the view that the status of mental phenomena is compromised if they are not reducible to something non-mental, seem to rest upon exactly the sort of *aprioristic* assumptions that naturalistic epistemology and philosophy of science have criticized. In point of fact, inter-theoretic reductions turn out to be quite rare in the sciences, and the popularity of reductionism in mid-20th-century philosophy of science was a product of the aprioristic and normative approach of the Logical Positivists and Empiricists. Attention to what one actually finds at the junctures between, say, physics and chemistry or biology suggests that

we should *not* take inter-theoretic reducibility as a normative principle in evaluating the special sciences generally. And if we should not adopt such a principle generally, there is little reason to think we should apply it just in the case of psychology. In short, empirical naturalism and *normative* philosophical naturalism make for *poor* bedfellows.

There is likewise a tension between empirical naturalism and philosophical projects that attempt to address questions that go beyond what the empirical sciences could possibly address. When one asks whether a regularity in nature is not only genuine and robust, but also *necessary*, one is often going beyond the bounds of the sciences. The empirical naturalist might find necessity claims innocuous when they are a result of genuine reductions or of applications of the principle of non-contradiction (that is, tautologies). But some problems in philosophy of mind are entangled in issues in modal metaphysics that are more problematic. If, for example, there are indeed truths that are metaphysically necessary, but in a way that is epistemically opaque to human minds, it would be impossible to differentiate them from merely nomic connections. And insofar as some debates in philosophy of mind trade upon, say, whether mind-body connections are contingent-but-nomic vs. necessary-but-inscrutable, these cannot be adjudicated through empirical means alone. If a specific brand of philosophical naturalism deems the answer to such a question to be important, she is committed to a problem that the empirical naturalist is likely to think we should avoid.

Here again I think it is useful to return to the historical pre-eminence of broad reductionism within philosophical naturalism. For philosophical naturalists, one of the reasons broad reductions have been so enduringly appealing is the fact that successful reductions carry with them implications of metaphysical supervenience. And indeed, aside from mathematical truths and tautologies, there is precious little that *does* provide the kind of decisive argument for metaphysical necessity that reductions offer. And so, to the extent that philosophical naturalism, as a project, is viewed as needing to offer metaphysically sufficient conditions for mental phenomena, its prospects seem closely tied to inter-theoretic reduction. (A philosophical naturalism that does not insist on metaphysically sufficient conditions will be compatible with dualism, a paradigmatically anti-naturalist position.) And since the prospects for inter-theoretic reduction seem slim indeed, philosophical naturalism would seem to be in deep trouble.

Empirical naturalism is in some ways much more promising. Insofar as it represents a commitment to studying the mind scientifically, there is little for anyone to object to in *that*. Even dualists from Descartes to Chalmers have vigorously embraced the scientific study of the mind. But insofar as empirical naturalism involves a further, *negative*, commitment to disavow any non-scientific form of knowledge or inquiry, it is more controversial. The things that the empirical sciences can reveal about the mind probably cannot, indeed, be gotten at by armchair argumentation,

and thus philosophy is no substitute for science. But this falls short of establishing that there are *no* non-empirical philosophical methods that might yield different types of understanding as well.

9. Conclusion

The term 'naturalism' is used to signify a number of philosophical positions in philosophy of mind. While most contemporary philosophers of mind claim to be naturalists, the variety of views styled 'naturalistic', combined with the fact that some of them are in fact incompatible with one another, makes it unlikely that there is any one such view that is anything like a consensus view today.

In its basic commitment to making use of the sciences of the mind to guide our philosophical understanding of the mind, empirical naturalism is committed to nothing that cannot be embraced by proponents of any of the mainstream views of the metaphysics of mind – reductive and non-reductive physicalists, eliminativists, dualists, or pluralists. (Indeed, it is acceptable to most Pragmatists, Social Constructionists and Idealists as well.) However, if the empirical naturalist goes beyond this, and holds that there are no truly philosophical problems of a trans-empirical nature, this is a view that would be rejected by many philosophers committed to traditional metaphysical projects, including both philosophical naturalists and non-naturalists.

Philosophical naturalism itself comes in a number of varieties, distinguished by whether they are claims about explanation or metaphysics, what style of scientific explanation they make use of, and whether they are presented as positive or normative claims. Normative forms of philosophical naturalism are in substantial tension with empirical naturalism. And among positive forms of philosophical naturalism, only reductive and non-reductive physicalism are truly incompatible with paradigmatically non-naturalistic commitments. Reductive physicalism is severely challenged by the turn away from reductionism in philosophy of science. Non-reductive physicalism is a consistent position, but its commitment to physicalism lacks the kind of strong evidence that would be supplied by reductions.

Short Biography

Steven Horst's research is principally in philosophy of mind and related topics in philosophy of cognitive science, metaphysics, epistemology and moral psychology. His *Symbols, Computation and Intentionality* (University of California Press, 1996) presents a critique of the Computational Theory of Mind based upon an analysis of the nature of symbolic meaning and computation. In his recent critique of naturalism, *Beyond Reduction* (Oxford University Press, 2007), Horst argues that contemporary philosophy of science presents enormous problems for all major views in philosophy

of mind, and develops an alternative called Cognitive Pluralism. *Laws, Mind and Freedom* (MIT Press, forthcoming) argues the sciences of the mind contain both laws and other types of rigorous models, and that libertarian freedom is compatible with a commitment to scientific laws, if the latter are understood in terms of causal capacities. He is now working on a book on Cognitive Pluralism. Educated at Boston University (B.A., 1982) and the University of Notre Dame (Ph.D., 1990), Horst is now Professor of Philosophy at Wesleyan University in Middletown, CT, where he has taught since 1990. He has also been a Visiting Scholar at the Center for Adaptive Systems at Boston University, the Center for the Study of Language and Information at Stanford, and Princeton University's Philosophy Department. His work has been recognized with two NEH Fellowships.

Notes

* Correspondence address: Department of Philosophy, Wesleyan University, Middletown, Connecticut, United States. Email: shorst@wesleyan.edu.

¹ There is, of course, a separate usage of the words 'naturalism' and 'naturalist' used to denote Aristotelian philosophy. This reflects an older Aristotelian use of the word 'nature(s)', which denoted the specific natures of particular types of things. 'Natural explanation', in Aristotelian/Scholastic parlance, was explanation that appealed to principles of change internal to the thing that was changing. For example, to explain why a spider builds a web or an acorn grows into an oak tree, one must appeal to the 'natures' of spiders and oaks, respectively.

² It is important to distinguish the contemporary use of the words 'mechanism' and 'mechanistic' from their use in the 17th century. The main difference is that in the older usage, 'mechanistic' explanations were restricted to those that proceeded through contact interactions, such as particle collisions or parts of a machine pushing against each other. No such restriction is implied in the contemporary usage.

³ Jaegwon Kim (*Supervenience and Mind*) argues against such independent mental-mental causation on grounds of double causation. David Papineau (*Philosophical Naturalism*) argues against it on the basis of a commitment to the causal closure of physics. But see Horst, *Beyond Reduction*.

Works Cited

- Allen, Colin and Marc Bekoff. *Species of Mind: The Philosophy and Biology of Cognitive Ethology*. Cambridge, MA: MIT Press, 1997.
- Baker, Lynne Rudder. *Explaining Attitudes*. Cambridge, UK: Cambridge UP, 1995.
- Barkow, Jerome H., Leda Cosmides and John Tooby. *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. New York, NY: Oxford UP, 1992.
- Bechtel, William and Adele Abrahamsen. 'Explanation: A Mechanist Alternative'. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 36.2 (2005): 421–41.
- Bechtel, William and Robert C. Richardson. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Princeton, NJ: Princeton UP, 1993.
- Bickle, John. *Philosophy and Neuroscience: A Ruthlessly Reductive Account*. Norwell, MA: Kluwer, 2003.
- . *Psychoneural Reduction: The New Wave*. Cambridge, MA: MIT Press, 1998.
- Botterill, George and Peter Carruthers. *The Philosophy of Psychology*. Cambridge: Cambridge UP, 1999.
- Brandom, Robert. *Making it Explicit*. Cambridge, MA: Harvard UP, 1994.

- . 'Modality, Normativity, and Intentionality'. *Philosophy and Phenomenological Research* 63.3 (2001): 587–609.
- Callebaut, Werner. *Taking the Naturalistic Turn, or, How Real Philosophy of Science Is Done*. Chicago, IL: U of Chicago P, 1993.
- Carnap, Rudolf. *Der Logische Aufbau der Welt*. Berlin: Weltkreis, 1928.
- . 'Psychology in Physical Language'. *Logical Positivism*. Ed. A. J. Ayer. New York, NY: Free Press, 1959. 165–98.
- Carruthers, Peter. *The Architecture of the Mind*. Oxford and New York, NY: Oxford UP, 2006.
- Chalmers, David. *The Conscious Mind: In Search of a Fundamental Theory*. New York, NY: Oxford UP, 1996.
- Chomsky, Noam. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press, 1965.
- . *Cartesian Linguistics*. New York, NY: Harper and Row, 1966.
- . *Language and Mind*. 3rd ed. Cambridge: Cambridge UP, 2006.
- Churchland, P. M. 'Eliminative Materialism and the Propositional Attitudes'. *Journal of Philosophy* 78 (1981): 67–90.
- Churchland, Patricia S. *Neurophilosophy: Towards a Unified Science of the Mind-Brain*. Cambridge, MA: MIT Press, 1986.
- Collier, John and Michael Stingl. 'Evolutionary Naturalism and the Objectivity of Morality'. *Biology and Philosophy* 8 (1993): 47–60.
- Cummins, Robert. "'How Does It Work?'" Versus "What Are the Laws?": Two Conceptions of Psychological Explanation'. *Explanation and Cognition*. Eds. Frank G. Keil and Robert Andrew Wilson. Cambridge, MA: MIT Press, 2000. 117–44.
- . *The Nature of Psychological Explanation*. Cambridge, MA: MIT Press, 1983.
- Davidson, Donald. 'Mental Events'. *Experience and Theory*. Eds. L. Foster and J. Swanson. New York, NY: Humanities Press, 1970. 79–102.
- Descartes, René. *The Philosophical Writings of Descartes*. Trans. J. Cottingham et al. 3 vols. Cambridge: Cambridge UP, 1985.
- Devitt, Michael. 'Naturalism and the a Priori'. *Philosophical Studies* 92 (1998): 46–65.
- and Kim Sterelny. *Language and Reality: An Introduction to the Philosophy of Language*. Cambridge, MA: MIT Press, 1999.
- Dretske, Fred. *Explaining Behavior*. Cambridge, MA: MIT Press, 1988.
- . 'If You Can't Make One, You Don't Know How It Works'. *Philosophical Naturalism. Vol. 19*. Eds. Peter A. French, Theodore E. Uehling, Jr. and Howard K. Wettstein. Notre Dame, IN: U of Notre Dame P, 2002. 468–82.
- . *Naturalizing the Mind*. Cambridge, MA: MIT Press, 1995.
- Fechner, Gustav Theodor. In *Sachen Der Psychophysik*. Leipzig: Breitkopf & Hartel, 1877.
- Fodor, Jerry. *Psychosemantics*. Cambridge, MA: Bradford Books, 1987.
- . *Representations*. Cambridge, MA: Bradford Books/MIT Press, 1981.
- . 'Special Sciences (Or: The Distunity of Science as a Working Hypothesis)'. *Synthese* 28 (1974): 97–115.
- . *A Theory of Content and Other Essays*. Cambridge, MA: Bradford Books, 1990.
- Garfinkel, Alan. 'Reductionism'. *The Philosophy of Science* [Chapter 2 from *Forms of Explanation*. New Haven, CT: Yale UP, 1981. 49–74]. Eds. Richard Boyd, Philip Gaspar and J. D. Trout. Cambridge, MA: MIT Press, 1981/1999. 443–59.
- Gibbard, Allan. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Cambridge, MA: Harvard UP, 1990.
- Godfrey-Smith, Peter. *Complexity and the Function of Mind in Nature*. Cambridge: Cambridge UP, 1996.
- Goldman, Alvin. *Readings in Philosophy and Cognitive Science*. Cambridge, MA: MIT Press, 1993.
- Hardcastle, Valerie. *The Myth of Pain*. Cambridge, MA: MIT Press, 1999.
- Hare, R. M. *The Language of Morals*. Oxford: Clarendon Press, 1952.
- Haugeland, John. *Mind Design II: Philosophy, Psychology, Artificial Intelligence*. Cambridge, MA: MIT Press, 1997.
- Helmholtz, Hermann von. *Handbuch Der Physiologischen Optik*. Leipzig: Voss, 1867.
- Hobbs, Jesse. 'Naturalism: A Contemporary Shibboleth?'. Unpublished manuscript presented at the NEH Summer Institute on Naturalism, held at the University of Nebraska, 1993.

- Horgan, Terry and John Tienson. 'The Intentionality of Phenomenology and the Phenomenology of Intentionality'. *Philosophy of Mind: Classical and Contemporary Readings*. Ed. D. J. Chalmers. New York, NY: Oxford UP, 2002. 520–32.
- Horst, Steven. *Beyond Reduction: Philosophy of Mind and Post-Reductionist Philosophy of Science*. New York, NY: Oxford UP, 2007.
- . 'Evolutionary Explanation and the Hard Problem of Consciousness'. *Journal of Consciousness Studies* 6.1 (1999): 39–48.
- . 'Phenomenology and Psychophysics'. *Phenomenology and the Cognitive Sciences* 4 (2005): 1–21.
- . *Symbols, Computation and Intentionality: A Critique of the Computational Theory of Mind*. Berkeley and Los Angeles, CA: University of California Press, 2007.
- Hume, David. *A Treatise on Human Nature*. Eds. David Fate Norton and Mary Norton. Oxford: Clarendon Press, 2000.
- Jackendoff, Ray S. *Consciousness and the Computational Mind*. Bradford Books. Cambridge, MA: MIT Press, 1987.
- Jackson, Frank. 'Epiphenomenal Qualia'. *Philosophical Quarterly* 32 (1982): 127–66.
- . *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Clarendon Press, 1998.
- . *Mind, Method and Conditionals*. New York, NY: Routledge, 1998.
- Kim, Jaegwon. 'Multiple Realization and the Metaphysics of Reduction'. *Philosophy and Phenomenological Research* 52 (1992): 1–26.
- . *Supervenience and Mind: Selected Philosophical Essays*. Cambridge: Cambridge UP, 1993.
- Kripke, Saul. *Naming and Necessity*. Cambridge, MA: Harvard UP, 1972.
- Kuhn, Thomas. *The Structure of Scientific Revolutions*. Chicago, IL: U of Chicago P, 1962.
- Lakoff, George. *Women, Fire and Dangerous Things: What Categories Reveal About the Mind*. Chicago, IL: U of Chicago P, 1987.
- and Mark Johnson. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York, NY: Basic Books, 1999.
- Laming, Donald. *The Measurement of Sensation*. New York, NY: Oxford UP, 1997.
- Levine, Joseph. 'Materialism and Qualia: The Explanatory Gap'. *Pacific Philosophical Quarterly* 64 (1983): 354–61.
- McGinn, Colin. *The Problem of Consciousness: Essays Towards a Resolution*. Oxford: Blackwell, 1991.
- Maturana, Humbert R. and Francisco Varela. *Autopoiesis and Cognition: The Realization of Living*. Dordrecht: Springer, 1980.
- Melnyk, Andrew. 'Can Physicalism be Non-Reductive?'. *Philosophy Compass* 3 (2008): 1281–96.
- Millikan, Ruth. *Language, Thought and Other Biological Categories*. Cambridge, MA: MIT Press, 1984.
- . *On Clear and Confused Ideas*. Cambridge: Cambridge UP, 2000.
- Moore, G. E. *Principia Ethica*. Cambridge: Cambridge UP, 1903.
- Nagel, Ernest. 'Naturalism Reconsidered'. *Proceedings and Addresses of the American Philosophical Association* 28 (1955): 5–17.
- . *The Structure of Science*. New York, NY: Harcourt, Brace and World, 1961.
- Nagel, Thomas. 'What Is It Like to Be a Bat?'. *Philosophical Review* 4 (1974): 435–50.
- Newton, Isaac. *Sir Isaac Newton's Mathematical Principles of Natural Philosophy and System of the World* (1713). Trans. Andrew Motte and Florian Cajori. Berkeley, CA: U of California P, 1962.
- Oppenheim, Paul and Hilary Putnam. 'Unity of Science as a Working Hypothesis'. *Concepts, Theories, and the Mind-Body Problem*. Vol. 2. Eds. H. Feigl, M. Scriven and G. Maxwell. Minnesota Studies in the Philosophy of Science. Minneapolis, MN: U of Minnesota P, 1958. 3–36.
- Papineau, David. *Philosophical Naturalism*. Oxford: Basil Blackwell, 1993.
- . 'Naturalism'. *The Stanford Encyclopedia of Philosophy (Spring 2007 Edition)*. Ed. Edward N. Zalta. 2007. <<http://plato.stanford.edu/archives/spr2007/entries/naturalism/>>.
- . *Thinking about Consciousness*. Oxford: Oxford UP, 2002.
- Penco, Carlo, Michael Beaney and Vignolo Massimiliano. *Explaining the Mind: Naturalist and Non-Naturalist Approaches to Mental Acts and Processes*. Cambridge: Cambridge Scholars Publishing, 2007.

- Pinker, Steven. *How the Mind Works*. New York, NY: Norton, 1997.
- Place, U. T. 'Is Consciousness a Brain Process?' *British Journal of Psychology* 47 (1956): 44–50.
- Polger, Thomas. *Natural Minds*. Cambridge, MA: MIT Press, 2004.
- Pollock, John and Robert Cummins. *Philosophy and Ai: Essays at the Interface*. Cambridge, MA: MIT Press, 1995.
- Port, Robert F. and Timothy Van Gelder. *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge, MA: MIT Press, 1995.
- Posner, Michael. *Foundations of Cognitive Science*. Cambridge, MA: MIT Press, 1989.
- Putnam, Hilary. 'Brains and Behavior'. *Readings in the Philosophy of Psychology*. Program of American Association for the Advancement of Science, Section L (History and Philosophy of Science), December 27, 1961. Ed. Ned Block. Cambridge, MA: Harvard UP, 1961/1980.
- . 'The Meaning of "Meaning"'. *Minnesota Studies in the Philosophy of Science* 7 (1975): 131–93.
- . 'Why Reason Can't be Naturalized'. *Synthese* 52 (1982): 3–24.
- Quine, W. V. O. 'Epistemology Naturalized'. *Ontological Relativity and Other Essays*. New York, NY: Columbia UP, 1969.
- . *Pursuit of Truth*. Cambridge, MA: Harvard UP, 1992.
- Ratliff, F. 'A Discourse on Edges'. *Visual Psychophysics and Physiology*. Eds. J. C. Armington, J. Krauskopf and B. R. Wooten. New York, NY: Academic Press, 1978.
- and L. Sirovich. 'Equivalence Classes of Visual Stimuli'. *Vision Research* 18 (1978): 845–51.
- Richards, Robert J. 'A Defense of Evolutionary Ethics'. *Biology and Philosophy* 1 (1986): 265–93.
- Rosenberg, Alex. 'A Field Guide to Recent Species of Naturalism'. *British Journal for the Philosophy of Science* 47.1 (1996): 1–29.
- Rouse, Joseph. 'New Philosophies of Science in North America – Twenty Years Later'. *Journal for the General Philosophy of Science* 29.1 (1998): 71–122.
- Ruse, Michael. *Evolutionary Naturalism: Selected Essays*. New York, NY: Routledge, 1995.
- . *Taking Darwin Seriously*. London: Basil Blackwell, 1986.
- Savage, C. Wade. *The Measurement of Sensation*. Berkeley, CA: U of California P, 1970.
- Searle, John R. *The Rediscovery of the Mind*. Cambridge, MA: MIT Press, 1992.
- Shapiro, Lawrence A. 'The Nature of Nature: Rethinking Naturalistic Theories of Intentionality'. *Philosophical Psychology* 10.3 (1997): 309–23.
- Siewert, Charles. *The Significance of Consciousness*. Princeton, NJ: Princeton UP, 1998.
- Silberstein, Michael. 'Reduction, Emergence and Explanation'. *The Blackwell Guide to the Philosophy of Science*. Eds. Peter Machamer and Michael Silberstein. Malden, MA and Oxford: Blackwell, 2002. 80–107.
- Smart, J. J. C. 'Sensations and Brain Processes'. *Philosophical Review* 68 (1959): 141–56.
- Sterelny, Kim. *The Representational Theory of Mind*. Oxford: Basil Blackwell, 1990.
- . *Thought in a Hostile World*. Oxford: Blackwell, 2003.
- Stich, Stephen P. *From Folk Psychology to Cognitive Science*. Cambridge, MA: MIT Press, 1983.
- . 'What is a Theory of Mental Representation?' *Mind* 101.402 (1992): 243–61.
- and Stephen Laurence. 'Intentionality and Naturalism'. *Midwest Studies in Philosophy*. Vol. 12. Eds. Peter A. French, Theodore E. Uehling, Jr. and Howard K. Wettstein. Notre Dame, IN: U of Notre Dame P, 1994. 159–82.
- Suppes, Patrick. 'The Structure of Theories and the Analysis of Data'. *The Structure of Scientific Theories*. Ed. Frederick Suppe. Urbana, IL: U of Illinois P, 1974.
- Todorovic, D. 'The Craik-O'Brien-Cornsweet Effect: New Varieties and Their Theoretical Implications'. *Perception & Psychophysics* 42 (1987): 545–60.
- Toulmin, Stephen. 'The Structure of Scientific Theories'. *The Structure of Scientific Theories*. Ed. Frederick Suppe. Urbana, IL: U of Illinois P, 1974.
- Tye, Michael. 'Naturalism and the Mental'. *Mind* 101.403 (1992): 421–41.
- Wagner, Steven J. and Richard Warner. *Naturalism: A Critical Appraisal*. South Bend, IN: U of Notre Dame P, 1993.
- Weber, Ernst Heinrich. *Du Pulsu, Resorptione, Audiu Et Tactu. Annotationes Anatomicae Et Physiologicae*. Leipzig: Koehler, 1834.
- Wright, Cory and William Bechtel. 'Mechanisms and Psychological Explanation'. *Handbook of Philosophy of Psychology and Cognitive Science*. Eds. Paul Thagard, Dov M. Gabbay and John Woods. Dordrecht: Elsevier, 2007. 31–79.