

## Editor's Introduction

D.M. WALSH

The papers collected in this volume are the proceedings of the 1999 Royal Institute of Philosophy conference: the theme of the conference, the same as the title of this collection, *Naturalism, Evolution and Mind*. The essays collected here cover a wide array of disparate themes in philosophy, psychology, evolutionary biology and the philosophy of science. They range in subject matter from the mind/body problem and the nature of philosophical naturalism, to the naturalization of psychological norms to the naturalization of phenomenal and intentional content, from the methodology cognitive ethology to issues in evolutionary psychology. They are united by the simple thought that the great promise of current naturalism in philosophy of mind resides in its potential to reveal mental phenomena as continuous with other phenomena of the natural world, particularly with other biological phenomena.

The study of the mind is undergoing something of a revolution. Advances in a number of disparate disciplines have given new impetus to the study of long-standing problems in the philosophy of mind. The essays collected in this volume give an overview of some of the ways in which developments in cognitive ethology, the philosophy of biology, evolutionary psychology, cognitive science and other fields have impacted upon the philosophy of mind.

Given the diversity of approaches and points of view on display it is something of a vain hope that a synoptic introduction might extract a single, overarching theme from this collection. To impose one would be artificial, not to say unilluminating. These papers may not present a single point of view or even cluster around a single, relatively well circumscribed issue, but that is not to say that they constitute only a gerrymandered collection. It is fascinating to see the way in which various disparate themes emerge, recur and intertwine among the papers: the significance of the concept of adaptation to the study of mind, the importance of higher-order intentionality, the role of the concept of function in naturalizing intentionality and phenomenal character, the significance of normativity to the understanding of mental phenomena. These themes and others appear repeatedly in the essays which follow.

The collection begins with an object lesson in the ways in which thinking of the mind as a biological category illuminates traditional

**D. M. Walsh**

philosophical problems. Larry Shapiro's *Mind The Adaptation* applies evolutionary thinking to the mind/body problem. Shapiro offers us a simple proposal: the mind is an adaptation. This proposal is contrasted with two existing conceptions of the mental: type-identity physicalism and functionalism. To its credit the mind-as-adaptation thesis steers a course between the 'baseless chauvinism' of type-identity physicalism and the promiscuity of functionalism. It is a commonplace that type-identity materialism constrains the domain of the mental too narrowly. If pain is identical to a certain kind of neurological state then creatures who are not endowed with the sort of physical machinery capable of realizing that state are also incapable of feeling pain. Functionalisms of various sorts have often been thought of as the antidote to the parochiality of type identity theory. According to functionalism, mental states are defined in terms of inputs and outputs *irrespective* of the sort of machinery that realizes them. But, so the usual rejoinder goes, if anything that realizes the set of inputs and outputs which characterizes a mental state *has* that mental state we may find ourselves, implausibly, counting systems like the economy of Bolivia and the nation of China among the bearers of mental states.

Shapiro argues that construing minds as adaptations captures our intuitions about the sorts of things capable of having minds with greater fidelity. Furthermore the concept of an adaptation may well serve to regiment or refine these largely pre-theoretic intuitions. The concept of an adaptation plays an especially significant part in modern systematics. We may find ourselves with a more tractable conception of the nature of mind if we are able to replace a set of largely pre-theoretic intuitions with one which has earned its keep within evolutionary biology.

An adaptation, according to Shapiro, 'is a structure plus a function', where the function of a biological trait is simply the effect that it has been selected for. Because being an adaptation of a certain sort requires that a trait has both a particular structure *and* a particular history, the individuation of mental categories is externalistic. Traits with the same structure may not count as being traits of the same kind. Contrary to physicalism then, sameness of physical realization is not sufficient for sameness of mental state. Of course, the usual objection to physicalism is that the conditions it sets on an entity's having a mind are already too restrictive. So far all we have from the mind-as-adaptation proposal is that these conditions are insufficient. Mind as adaptation has yet to demonstrate that they are unnecessary too. After all, the intuition that type physicalism violates is that mind ought to accommodate at least some form of

## Editor's Introduction

multiple realizability. This issue is addressed in Shapiro's comparison of functionalism with the mind-as-adaptation view.

The problem with functionalism seems to be that it opens up the prospect of untrammelled multiple realizability. 'It is only when one chooses to individuate kinds on the basis of function alone that one can remain neutral about the nature of the stuff that fills the functional role. However, it is just this neutrality that also seems to reduce functionalism to absurdity...' (p. 33) Mind-as-adaptation permits a certain amount of multiple realizability—anything that is structured like a brain can be a mind so long as it has the right selectional (that is *not* to say, phylogenetic) history. But how much multiple realizability is allowed? Shapiro here invokes a principle from functional morphology: that function constrains structure. There are only a certain few ways to build a functioning eye. As it turns out, traits which have been selected independently for subserving the function of eyes show remarkable structural similarities. Shapiro conjectures that the same sort of constraints apply with minds. The intuitions upon which claims of the wild multiple realizability of minds are based are in conflict with empirical facts about the ways in which function constrains structure. Something which subserves the function of a mind could not be structured like whatever it is that physically realizes the economy of Bolivia.

'Mind-as-adaptation is...shorthand. It's a slogan that invites us to conceive of the various cognitive capacities that we typically associate with the mind as structures with functions' (Shapiro p. 36). This issue arises again in later essays. Shapiro clearly presents rationale for the evolutionary psychology discussed by Wheeler and Atkinson and Plotkin. It is also interesting to note the tension between Shapiro's justification for the use of history in individuating mental categories and Jackson's claim that the mental categories with which we are most familiar are functional and not historical.

Thomas Bontly's (*Should Intentionality be Naturalized?*) discussion of naturalism raises another set of issues related to the realization of mental states by physical states. Naturalism is both a constraint upon, and a goal of, much of current philosophy of mind, but Bontly's discussion elegantly reminds us that few if any of us are able to say what naturalism is. For some, naturalism is simply a methodological thesis (a position reflected in Sober's contribution)—to be natural is to subject to investigation by a battery of methods recognizable as scientific. Bontly's claim is that naturalism, if it is a substantive thesis at all, is a thesis about the way in which mental states are realized in the world described by physics. After

**D. M. Walsh**

all, a central tenet of naturalism seems to be that the respective ontologies of the various branches of science comprise a unity.

One common way to cash out the presumed unity implicit in naturalism is to appeal to causal closure. Everything that is natural, it is commonly supposed, comprises a domain which is causally closed and complete. To make a case that putative entities of a certain kind constitute part of the natural world one simply has to show their role in the causal structure of the world. The way to demonstrate that intentional states are natural is to show how they are realized in the world by states susceptible of strictly physical description. The way to demonstrate the autonomy of the intentional sciences seems to require that intentional states play some irreducible causal role in the closed causal structure of the physical world. But the rub is that the causal closure of the physical world is incompatible with the causal efficacy of mental states. If we insist that the mental must be realized in the physical we risk either the causal exclusion of the mental *or* causal overdetermination in the sense that every event with a mental cause also has a sufficient physical cause.

Bontly is dubious of this sort exclusion argument. It threatens to introduce either exclusion or overdetermination for every cause which supervenes on some physical state or other. Supervenient properties are not in general excluded by their subvening physical realizers. So the fact that mental states supervene on physical states should not exclude them from being independently causally efficacious. Nor, in general, is it the case that demonstrating the causal efficacy of a property of a given kind requires demonstrating the manner of its physical realization. There is little reason to suppose that the overly stringent naturalistic standards that have been applied in the realm of intentional psychology are any more appropriate there than in any other field of the natural sciences.

After the discussion of mind/body matters in the first two papers, attention turns to three long-standing problems posed to naturalism by the nature of mental states: phenomenal content, normativity and intentionality. In each of these cases, recent work in higher-order intentionality and teleosemantics offer distinctive, sometimes conflicting perspectives on these issues. Peter Carruthers (*Consciousness: explaining the phenomena*) asks whether phenomenal consciousness is susceptible of reductive, naturalistic explanation. It is thought by many that such a reductive project is bootless on account of the distinctive feature of phenomenal consciousness which puts it beyond the ambit of naturalism. Phenomenal conscious states are, well, phenomenal; there is something it is like to undergo them. This what it is like, it seems, is ineffable, at least it

**Editor's Introduction**

evades capture in the descriptive physicalist vocabulary. Phenomenal states are also, evidently, subjective. These features open an 'explanatory gap' which the physical sciences evidently cannot breach.

Carruthers acknowledges a significant aspect of the explanatory gap; we have a capacity to recognize phenomenal properties of experience 'non-inferentially or "straight-off"'. Our concepts of what constitute phenomenal experience are bare recognitional ones. He demurs at the claim that the phenomenal properties by which we recognize phenomenal states must be, as many mysterians insist, *intrinsic* properties. According to Carruthers, what needs to be *explained* by the naturalist is the *subjective feel* of phenomenal states; what needs to be explained *away* is the apparent intrinsicness of their properties. The explanatory gap exists because of competing ways in which we conceptualize the mental, from the scientific or introspective perspectives, but the fact that we have distinctive ways of conceptualizing the phenomenal doesn't entail that these conceptualizations do not pick out the same features of the world. Naturalizing phenomenal consciousness involves giving a scientifically acceptable account of those features of the natural world that fall under our bare recognitional concepts.

Carruthers offers a dispositionalist Higher Order Thought theory of phenomenal consciousness. In Carruthers' scheme conscious experiences are '*analog intentional contents*' (p. 70). They are perceptual states whose contents represent states of the world, which in turn are apt (disposed) to form the contents of higher-order thoughts. The contents of phenomenal states are 'analog' (or analog-like) in virtue of the fact they are more fine-grained than any of our descriptive concepts. The content of phenomenal experiences outstrips our capacity to subsume it entirely under the battery of descriptive concepts at our disposal—hence the famous ineffability of phenomenal experience.

Controversially, Carruthers supposes that the Higher Order Thoughts which constitute phenomenal experiences require a full-blown theory of mind system sufficiently sophisticated to understand the *is/seems* distinction. This level of sophistication, it is thought is generally not attained in children until the age of at least 3.

The dispositionalist HOT theory can explain away the apparent existence of qualia (the distinctive, apparently intrinsic properties of experience). Perceptual states represent objects as having intrinsic properties (Carruthers uses the example *redness*). The higher order *seemings* of redness, will then be '*... seemings of the presence of a certain intrinsic property*' (p. 75). Yet the contents of these states will be representational and not intrinsic. The

**D. M. Walsh**

subjectivity of phenomenal feel can also be accounted for. It is here that Carruthers appeals to teleosemantics. According to consumer semantics variants of teleosemantics, the representational content of a state depends upon the consumer systems which use that state. Because conscious perceptual states are used by the mind-reading HOT device, the contents of these states have a dual-nature. They represent the world to us as being in such and such a state *and* by dint of being the contents of HOTs represent *that* the world is being perceived by us to be in such and such a state: ‘... each perceptual state with the content *red* is already a state with the higher-order analog content *seems red* or *experience of red*. Each such state then has, as part of its content, a dimension of *seeming* or *subjectivity*.’ (p. 76).

Carruthers points to the significant role of consumer semantics in naturalizing the phenomenal content of experience. The theme of consumer semantics surfaces in each of the following two papers. It is implicitly denied in Dretske’s account of the respective roles of history and norms in naturalizing intentionality. It is explicitly recommended in Millikan’s discussion of the significance of natural information to the nature of intentional content.

Dretske (*Norms, History and the Mental*) addresses a problem seen by many as the scourge of naturalism, the norm-ladenness of the mental. Mental concepts, it is often supposed, are imbued with normativity. Normativity enters into considerations of the mental in a number of ways. Psychological states—beliefs and desires—are individuated, it is often supposed, by their norm-governed roles in cognitive affairs. The goal of beliefs and judgments is truth. As such psychological states are subject to normative evaluation. There is something undesirable, or to be avoided, about false beliefs. If the natural world is norm-free then naturalism cannot sustain an account of mental states and the norm-governed relations they enter into. One common way to naturalize the normativity inherent in the mental is to suppose that norms are introduced by selectional (or learning) history—a trait is supposed to do *f* if it has been selected to do *f*. If the mental is essentially normative, and norms come from history, then the mental is essentially historical. Thoughts, like fossils, require a certain kind of history.

Dretske is dubious of the attempt to ground the normativity inherent in cognition in the putative norms of biology. Rather he simply repudiates the entire problem of naturalizing the norms of the mental. Neither the concepts of mental states, like *belief* and *judgment* nor the truth they aim at are norm-laden, argues Dretske. Truth and falsity are not normative categories: truth is merely

**Editor's Introduction**

correspondence. Nor is belief a normative category. It is not part of the essence of belief to aim at truth; a belief which serves its purpose only when it is false is no less a belief.

Dretske uses the analogy of a road sign. 'An arrow ... can point to Chicago or away from Chicago. There is a difference here, yes, but the difference is not normative. Aside from our purpose in putting the sign there or in using it as a guide..., there is nothing right or wrong, nothing that is supposed-to-be or supposed-not-to-be, about an arrow pointing to Chicago' (p. 92). The normativity often thought to be inherent in beliefs, truth, meanings etc. resides not in these things but in the intentions we have in deploying them. We avoid false beliefs because, by and large, they do not serve our purposes. 'The normativity of perceptual beliefs and experiences thus derives from the teleofunctions of the mechanisms producing them, and these teleofunctions emerge from the historical (selectional) process that shaped those mechanisms' (p. 98). But, Dretske goes on to argue, this isn't sufficient to establish the point that perceptual concepts are norm-laden. The norms inherent in our purposes are historical ones. Dretske concludes that '... although we may need history to explain norms, we do not need history to explain mentality because mentality, though subject to norms..., isn't constituted by them' (p. 99).

But, Dretske surmises, normativity may work its way into the determination of the mental by another route. If intentionality—the aboutness of mental states—is essentially historical, and norms devolve from history, intentionality may also be essentially norm-laden. Dretske resists the conclusion; intentional contents are history-laden, but they are not, on account of that, norm-laden.

Dretske's themes are taken up in various other places in this collection. Alan Millar presents a strikingly different account of the ways in which normativity might be thought to be integral to the mental. Furthermore, Dretske's distinctive conception of the roles of history and natural information in the determination of intentional content present the ideal platform for the discussions of teleosemantics offered by Millikan and Jackson which follow. Millikan challenges the notion of natural information implicit in Dretske's account of content. Jackson addresses the presumed significance of history. Millikan and Jackson do not however offer a unified front. Indeed they appear to take diametrically opposed views on an issue of central importance. Millikan believes that it is a requirement on any naturalized account of content that it countenance the possibility that a creature may have a concept that applies to *Xs* and not *Ys* even though that creature cannot distinguish

**D. M. Walsh**

between *Xs* and *Ys*. Jackson, in contrast, cites with approval, Locke's claim that 'Words being voluntary signs, they cannot be voluntary signs imposed on [a man] by things he knows not' (quoted in Jackson p. 127). These two positions look to be in tension.

Millikan (*What Has Natural Information to do with Intentional Representation?*) outlines the conception of 'natural information' which figures in many informational accounts of intentional content, most particularly those advocated by Dretske. According to Dretske, for an intentional state to represent a state of affairs that *s* is *F*, it must carry information that *s* is *F*. The relation between a signal, *r*, and the information it carries is cashed out in terms of probabilities. 'A signal *r* carries the information that *s* is *F* = The conditional probability of *s*'s being *F*, given *r* (and [background knowledge] *k* is 1 (but, given *k* alone, less than 1)' (quoted in Millikan p. 105). The probability of 1 between the signal and the occurrence of the state of affairs about which it carries information here is explicitly intended to convey the idea that there is a law of nature which determines that *r* is tokened only when *s* is *F*. Of course a signal *r* is capable of carrying information of the correct sort only if certain enabling conditions are met. Dretske calls these channel conditions. So, by the definition of natural information, *r* carries natural information that *s* is *F* only if the conditional probability that *s* is *F* given *r* is 1, given that a set of appropriate channel conditions obtains and that this probability is underwritten by natural law. The official definition of information makes no mention of the frequency with which the channel conditions are met. Millikan dubs this form of information 'InformationL'

Millikan asks what good it does for an animal to receive informationL about its environment. Here she diagnoses a problem for accounts of intentionality based on informationL. 'InformationL depends upon a channel between the information source and the signal producing a correspondence between the two in accordance with natural necessity. But unfortunately, relatively few things that an animal needs to know about figure as sources for this kind of information.' (p. 114). A mouse may need to know of the presence of a hawk overhead. However, it is not a matter of natural necessity that only hawks produce the sort of effects by which mice detect them. Millikan concludes that '[I]t is the *properties* of hawks...that enter into natural laws, not the hawks ... themselves, and it is never a matter of natural law that only hawks ... have these properties.' (p. 115).

Millikan asks whether some corrective can be applied to the notion of natural information to suit the purposes of a naturalized



**Editor's Introduction**

theory of intentional representation. It is a requirement on an adequate theory of intentional representation that it demonstrate how representations can be false or represent non-existent entities. Dretske's solution to this challenge, according to Millikan, is to overlay a theory of teleological function onto a conception of natural information. Indeed we saw this strategy at work in the preceding chapter. The representational content of a state is determined by the natural information it carries. The possibility of error is introduced by the claim that the apparatus in question—the animal's cognitive system—has been designed by evolution to convey informationL and sometimes that apparatus fails to operate as it was designed to do. The problem here is that it introduces the error at the wrong place. Most cases of misrepresentation are not the result of breakdown of the animal's perceptual or cognitive systems. They are the result of uncooperative environment. Nothing has gone wrong, for example, with an organism which represents its own reflection in a mirror as a potential mate. It simply finds itself in an environment which systematically undermines the design of its perceptual system.

Millikan insists that error is introduced in the wrong place because teleology has been introduced in the wrong place. She recommends, to borrow Carruthers' phrase, 'a dose of consumer semantics'. What determines whether a signal represents a state of affairs is not some nomic relation between the signal and the state of affairs but how the signal is *used* by a downstream consumer device. What determines how the signal is used, in turn, is fixed by the conditions under which the downstream device has produced outputs which, in the past, have been selectively beneficial. To be sure, there must be some correlation between the state represented (that *s* is *F*) and the representation, *r*, but this correlation does not need to be nomic covariation. So long as the signal *r* correlates sufficiently closely to the conditions under which the outputs of the consumer device are beneficial (*s*'s being *F*) then *r* can be said to represent that *s* is *F*. In these circumstance *r* carries information, of a sort, that *s* is *F*. Millikan calls this informationC. InformationC differs from informationL insofar as the latter but not the former involves conditional probabilities of 1. If informationC is to count as a form of natural information, it follows that a signal *r* could carry information about, that is to say represent, *X* even if *r* does, or would, also correlate with *Y*. Even if a creature could not discriminate *X*'s from *Y*'s it may be capable of having a representation of *X* which is not also a representation of *Y*. Millikan's contention is that only a 'softer' notion of information, such as informationC, resists the verifica-

**D. M. Walsh**

tionist implication that creature cannot have a representation whose content is  $X$  and not  $Y$  if it cannot distinguish  $X$ 's from  $Y$ s.

It is another question whether the verificationist implication is one that should be resisted. This question is usually answered in the affirmative. But Jackson isn't so sure. Jackson (*Locke-ing onto Content*) cites Locke to the effect that '...language ... rests on voluntary, if largely implicit, agreements to use words and sentences to stand for how we take things to be.' Jackson contends that this Lockean point raises a significant problem for precisely the sort of teleological account of semantics that Millikan endorses. Teleological theories of content typically rely on history. Jackson, citing Papineau, gives a rough sketch: '...the belief that  $P$  is the state selected to co-vary with  $P$ , and the desire that  $P$  is the state selected to bring about  $P$ ' (p. 130). But, one does not need to know the history of a particular intentional state in order to know its content. When we communicate the contents of our beliefs to one another, we communicate something about the way we take the world to be; we typically do not communicate the historical determinants of the state. We exploit what Jackson calls the 'folk' or 'functional' roles of our beliefs. There are two aspects to intentional content: current functional role, which is easily accessible and selectional history which isn't. Jackson argues that 'Locke's point tells us that the way we acquired language and its evident utility depend upon the contents of very many of our beliefs and desires being pretty much common knowledge; but if content is given by selectional history, these contents will be very far from common knowledge' (p. 131).

The teleosemanticist can choose either of two options. The first is to acknowledge that intentional states have two kinds of content, functional role (informational) content and selectional content. The second is to propose that specification of a state's history and its current functional role are merely two ways of latching onto the same intentional content. Jackson does not recommend the former. It makes teleosemantics a much less interesting theory. Teleosemantics is, after all, proffered as an account of *the* content of intentional states. The second option is more congenial to the teleosemanticist. It holds that intentional content is selectional content and that in turn is for the most part also consonant with current functional role. So to understand what the selectional content of an intentional state is, you only need to look at its current role. In much the same way, functional anatomists latch onto the function of an anatomical trait by studying its current functional role and this procedure is justified even if function is determined by selectional history.

But even this strategy has its disadvantages. Jackson points out