Chapter 11

Individualism

Robert A. Wilson

11.1 Introduction

Much discussion has been generated in the philosophy of mind over the last 25 years or so on the general issue of the relationship between the nature of the mind of the individual and the character of the world in which that individual, and hence her mind, exists. The basic issue here is sometimes glossed in terms of whether psychological or mental states are "in the head," but to the uninitiated that is likely to sound like a puzzling issue to debate: of course mental states are in the head! (but see Rowlands 1999; Wilson 2000a, 2001). So one of our first tasks is to articulate a version of the issue that makes more perspicuous why it has been a topic of some contention for so long, and that begins to convey something of its importance for a range of diverse issues, such as the methodology of cognitive science, the possibility of self-knowledge, and the nature of intentional representation.

Consider the question of whether the character of an agent's environment plays some crucial role in determining or fixing the nature of that agent's mind. A natural thought, one shared by those who disagree about the answer to the question above, would be that agents causally interact with their world, gathering information about it through their senses, and so the nature of their minds, in particular what their thoughts are about, *are* in part determined by the character of their world. Thus, the world is a *causal* determinant of one's thoughts, and thus one's mind. That is, the world is a contributing cause to the *content* of one's mind, to what one perceives and thinks about. This is just to say that the content of one's mind is not causally isolated from one's environment. Separating individualists and anti-individualists in the philosophy of mind is the question of whether there is some *deeper* sense in which the nature of the mind is determined by the character of the individual's world.

We can approach this issue by extending the brief discussion above of the idea that the content of the mind is in part causally determined by the agent's

environment to explore the conditions under which a difference in the world implies a difference in the mind. Individualists hold that this is so just in case that difference in the world makes some corresponding change to what occurs inside the boundary of the individual; anti-individualists deny this, thus allowing for the possibility that individuals who are identical with respect to all of their *intrinsic* features could nonetheless have psychological or mental states with different contents. And, assuming that mental states with different contents are *ipso facto* different types or kinds of state, this implies that an individual's intrinsic properties do not determine or fix that individual's mental states.

This provides us with another way, a more precise way, of specifying the difference between individualism (or *internalism*) and its denial, anti-individualism (or *externalism*), about the mind. Individualists claim, and externalists deny, that what occurs inside the boundary of an individual *metaphysically* determines the nature of that individual's mental states. The individualistic determination thesis, unlike the causal determination thesis, expresses a view about the nature or essence of mental states, and points to a way in which, despite their causal determination by states of the world, mental states are autonomous or independent of the character of the world beyond the individual. What individualism implies is that two individuals who are identical in all their intrinsic respects *must have* the same psychological states. This implication, and indeed the debate over individualism, is often made more vivid through the fantasy of *doppelgangers*, molecule-for-molecule identical individuals, and the corresponding fantasy of Twin Earth. I turn to these dual fantasies via a sketch of the history of the debate over individualism.

11.2 Getting to Twin Earth: What's in the Head?

Hilary Putnam's "The Meaning of 'Meaning'" (1975) introduced both fantasies in the context of a discussion of the meaning of natural language terms. Putnam was concerned to show that "meaning" does not and cannot jointly satisfy two theses that it was often taken to satisfy by then prevalent views of natural language reference: the claim that the meaning of a term is what determines its reference, and the claim that meanings are "in the head," where this phrase should be understood as making a claim of the type identified above about the metaphysical determination of meanings. These theses typified *descriptive* theories of reference, prominent since Frege and Russell first formulated them, according to which the reference of a term is fixed or metaphysically determined by the descriptions that a speaker attaches to that term. To take a classic example, suppose that I think of Aristotle as a great, dead philosopher who wrote a number of important philosophical works, such as the *Nicomachean Ethics*, and who was a student of Plato and teacher of Alexander the Great. Then, on a descriptivist view of reference, the reference of my term "Aristotle" is just the

thing in the world that satisfies the various descriptions that I attach to that term: it is the thing in the world that is a great philosopher, is dead, wrote a number of important philosophical works (e.g., *Nicomachean Ethics*), was a student of Plato, and was a teacher of Alexander the Great.

Such descriptivist views of the reference of *proper names* were the critical focus of Saul Kripke's influential *Naming and Necessity* (1980), while in his attack on this cluster of views and their presuppositions, Putnam focused on *natural kind terms*, such as "water" and "tiger." Both Kripke and Putnam intended their critiques and the subsequent alternative theory of natural language reference, the causal theory of reference, to be quite general and to provide an alternative way to think about the relationship between language and the world. But let us stay close to Putnam's argument and draw out its connection to individualism.

Consider an ordinary individual, Oscar, who lives on Earth and interacts with water in the ways that most of us do: he drinks it, washes with it, and sees it falling from the sky as rain. Oscar, who has no special chemical knowledge about the nature of water, will associate a range of descriptions with his term "water": it is a liquid that one can drink, that is used to wash, and that falls from the sky as rain. On a descriptive view of reference, these descriptions determine the reference of Oscar's term "water." That is, Oscar's term "water" refers to whatever it is in the world that satisfies the set of descriptions he attaches to the term. And since those descriptions are "in the head," natural language reference on this view is individualistic.

But now, to continue Putnam's argument, imagine a molecule-for-molecule doppelganger of Oscar, Oscar*, who lives on a planet just like Earth in all respects but one: the substance that people drink, wash with, and see falling from the sky is *not* water (i.e., H_2O), but a substance with a different chemical structure, XYZ. Call this planet "Twin Earth." This substance, whose chemical composition we might denote with "XYZ," is called "water" on Twin Earth, and Oscar*, as a doppelganger or twin of Oscar, has the same beliefs about it as Oscar has about water on Earth. (Recall that Oscar, and thus Oscar* as his twin, have no special knowledge of the chemical structure of water.) Twin Earth has what we might call "*twin-water*" or "*twater*" on it, not water, and it is twater that Oscar* interacts with, not water – after all, there is no water on Twin Earth. Given that Oscar's term "water" refers to or is about water, then Oscar*'s term "water" refers to or is about *twater*. That is, they have natural language terms that *differ* in their meaning, assuming that reference is at least one aspect of meaning. But, by hypothesis, Oscar and Oscar* are doppelgangers, and so are identical in all their intrinsic properties, and so are identical with respect to what's "in the head." Thus, Putnam argues, the meaning of the natural language terms that Oscar uses are not metaphysically determined by what is in Oscar's head.

Putnam's target was a tradition of thinking about language which was, in terms that Putnam appropriated from Rudolph Carnap's *The Logical Construction of the World* (1928), *methodologically solipsistic*: it treated the meanings of natural language terms and language more generally in ways that did not suppose that the

world beyond the individual language user exists. Since Putnam's chief point was one about natural language terms and the relationship of their semantics to what's inside the head, one needs at least to extend his reasoning from language to thought to arrive at a position that denies individualism about the mind itself. But given the tradition to which he was opposed, such an extension might be thought to be relatively trivial, since in effect those in the tradition of methodological solipsism – from Brentano, to Russell, to Husserl, to Carnap – conceived of natural languages and their use in psychological terms.

The introduction of the term "individualism" itself can be found in Tyler Burge's "Individualism and the Mental" (1979), where Burge developed a series of thought experiments in many ways parallel to Putnam's Twin Earth thought experiment. Burge identified individualism as an overall conception of the mind prevalent in modern philosophical thinking, at least since Descartes in the mid-seventeenth century, and argued that our common-sense psychological framework for explaining behavior, our *folk psychology*, was not individualistic. Importantly, Burge was explicit in making a case against individualism that did not turn on perhaps controversial claims about the semantics of natural kind terms – he developed his case against individualism using agents with thoughts about arthritis, sofas, and contracts – and so his argument did not presuppose any type of scientific essentialism about natural kinds. Like Putnam's argument, however, Burge's argument does presuppose some views about natural language understanding.

The most central of these is that we can and do have *incomplete* understanding of many of the things that we have thoughts about and for which we have natural language terms. Given that, it is possible for an individual to have thoughts that turn on this incomplete understanding, such as the thought that one has arthritis in one's thigh muscle. Arthritis is a disease only of the joints, or as we might put it, "arthritis" in our speech community applies only to a disease of the joints. Consider an individual, Bert, with the belief that he would express by saying "I have arthritis in my thigh." In the actual world, this is a belief about arthritis; it is just that Bert has an incomplete or partially mistaken view of the nature of arthritis, and so expresses a false belief with the corresponding sentence.

But now imagine Bert as living in a different speech community, one in which the term "arthritis" *does* apply to a disease of both the joints and of other parts of the body, including the thigh. In *that* speech community, Bert's thought would not involve the sort of incomplete understanding that it involves in the actual world; in fact, his thought in such a world would be *true*. Given the differences in the two speech communities, it seems that an individual with thoughts about what he calls "arthritis" will have different thoughts in the two communities: in the actual world, Bert has thoughts about arthritis, while in the counterfactual world he has thoughts about *some other disease* – what *we* might refer to as "tharthritis," to distinguish it from the disease that we have in the actual world. In principle, we could suppose that Bert himself is identical across the two contexts – that is, he is identical in all intrinsic respects. Yet we attribute thoughts with different contents to Bert, and seem to do so *solely* because of the differences

in the language community in which he is located. Thus, the content of one's thoughts is not metaphysically determined by the intrinsic properties of the individual. And again taking a difference in the content of two thoughts to imply a difference between the thoughts themselves, this implies that thoughts are not individuated individualistically.

One contrast that is sometimes (e.g., Segal 2000: chs. 2–3) drawn between the anti-individualistic views of Putnam and Burge is to characterize Putnam's view as a form of *physical* externalism and Burge's view as a form of *social* externalism: according to Putnam, it is the character of the physical world (e.g., the nature of water itself) that, in part, metaphysically determines the content of one's mind, while according to Burge it is the character of the social world (e.g., the nature of one's linguistic community) that does so. While this difference may serve as a useful reminder of one way in which these two views differ, we should also keep in mind the "social" aspect to Putnam's view of natural language as well: his linguistic division of labor. Important to both views is the idea that language users and psychological beings depend and rely on one another in ways that are reflected in our everyday, common-sense ways of thinking about language and thought. Thus there is a social aspect to the nature of meaning and thought on both views, and this is in part what justifies the appropriateness of the label anti-*individualism* for each of them.

11.3 The Cognitive Science Gesture

Philosophers who see themselves as contributing to cognitive science have occupied the most active arena in which the debate between individualists and externalists has been played out. At around the time that individualism was coming under attack from Putnam and Burge, it was also being defended as a view of the mind particularly apt for a genuinely scientific approach to understanding the mind, especially of the type being articulated within the nascent interdisciplinary field of cognitive science. For those offering this defense, there was something suspiciously unnaturalistic about the Putnam–Burge arguments, as well as something about their conclusions that seemed anti-scientific, and part of the defense of individualism and the corresponding attack on externalism turned on what I will call the *cognitive science gesture*: the claim that, as contemporary empirical work on cognition indicated, any truly scientific understanding of the mind would need to be individualistic.

Picking up on Putnam's use of "methodological solipsism", Jerry Fodor defended methodological solipsism as the doctrine that psychology ought to concern itself only with *narrow* psychological states, where these are states that do not presuppose "the existence of any individual other than the subject to whom that state is ascribed" (Fodor 1980: 244). Fodor saw methodological solipsism as the preferred way to think of psychological states, given especially the Chomskyan

revolution in linguistics and the accompanying computational revolution in psychology. If mental states were transitions governed by computational rules, then the task of the cognitive sciences would be to specify those rules; insofar as mental states were computational, broader considerations about the physical or social worlds in which an individual is located seem irrelevant to that individual's psychological nature.

Stephen Stich's (1978) principle of autonomy provides an alternative way to articulate an individualistic view of cognitive science, variations on which have become the standard ways to formulate individualism. The principle says that "the states and processes that ought to be of concern to the psychologist are those that supervene on the current, internal, physical state of the organism" (Stich 1983: 164–5). The notion of *supervenience* provides a more precise way to specify the type of metaphysical determination that we introduced earlier. A set of properties, S (the supervening properties), supervenes on some other set of properties, B (the base properties), just if anything that is identical with respect to the B properties must also be identical with respect to the S properties. In part because of the prominence of supervenience in formulating versions of physicalism, together with the perceived link between physicalism and individualism (more of which in a moment), but also in part because of the emphasis on doppelgangers in the Putnam and Burge arguments, it has become most typical to express individualism and its denial in terms of one or another supervenience formulation.

Common to both Fodor and Stich's views of cognitive science is the idea that an individual's psychological states should be *bracketed off* from the mere, beyond-the-head environments that individuals find themselves in. Unlike Putnam and Burge in the papers discussed above, Fodor and Stich have focused on the relevance of individualism for explanatory practice in psychology, using their respective principles to argue for substantive conclusions about the scope and methodology of psychology and the cognitive sciences. Fodor contrasted a solipsistic psychology with what he called a naturalistic psychology, arguing that since the latter (amongst which he included J. J. Gibson's approach to perception, learning theory, and the naturalism of William James) was unlikely to prove a reliable research strategy in psychology, methodological solipsism provided the only fruitful research strategy for understanding cognition (see also Fodor 1987). Stich argued for a syntactic or computational theory of mind which made no essential use of the notion of intentionality or mental content at all, and so used the principle of autonomy in defense of an *eliminativist* view about content (see also Stich 1983).

Although I think that the cognitive science gesture *is* a gesture (rather than a solid argument that appeals to empirical practice), it is not an *empty* gesture. While Fodor's and Stich's arguments have not won widespread acceptance in either the philosophical or cognitive science communities, they have struck a chord with those working in cognitive science, perhaps not surprisingly since the dominant research traditions in cognitive science have been at least implicitly individualistic. Relatively explicit statements of a commitment to an individualistic view of

aspects of cognitive science include Chomsky's (1986, 1995, 2000) deployment of the distinction between two conceptions of language (the "I"-language and the "E"-language, for "internal" and "external", respectively), Jackendoff's (1991) related, general distinction between "psychological" and "philosophical" conceptions of the mind, and Cosmides and Tooby's (1994) emphasis on the constructive nature of our internal, evolutionary-specialized cognitive modules.

Part of the attraction of individualism for practicing cognitive scientists is its perceived connection to the representational theory of mind, which holds that we interact with the world perceptually and behaviorally through internal mental representations of how the world is (as the effects of perceiving) or how the world should be (as instructions to act). Jackendoff expresses such a view when he says:

> Whatever the nature of real reality, the way reality can look to us is determined and constrained by the nature of our internal mental representations. . . . Physical stimuli (photons, sound waves, pressure on the skin, chemicals in the air, etc.) act mechanically on sensory neurons. The sensory neurons, acting as transducers in Pylyshyn's (1984) sense, set up peripheral levels of representation such as retinal arrays and whatever acoustic analysis the ear derives. In turn, the peripheral representations stimulate the construction of more central levels of representation, leading eventually to the construction of representations in central formats such as the 3D level model. (1991: 159–61)

Provided that the appropriate, internal, representational states of the organism remain fixed, the organism's more peripheral causal involvement with its environment is irrelevant to cognition, since the only way in which such causal involvement can matter to cognition is by altering the internal mental states that represent that environment.

11.4 Functionalism, Physicalism, and Individualism

For many philosophers interested in the cognitive sciences, individualism has been attractive because of a perceived connection between that view and both physicalism and functionalism in the philosophy of mind, both of which have been widely accepted since the 1980s. Physicalism (or materialism) is a view that has been expressed in various ways, perhaps the most common of which is in terms of the notion of supervenience: all facts, properties, processes, events, and things supervene on the physical facts, properties, processes, events, and things, as they are posited in elementary physics. This ontological formulation of physicalism (concerned with what *exists*) is often accompanied by an *explanatory* thesis, which states that physical explanations are, in some sense, the ultimate explanations for any phenomenon whatsoever.

Individualism has been thought to be linked to physicalism, since it implies, via the supervenience formulation, that there is no psychological difference without a corresponding difference in the intrinsic, physical states of the individual. Those rejecting individualism have sometimes been charged (e.g., by Block 1986 and Fodor 1987: ch. 2) with endorsing some form of dualism about the mind, or making a mystery of mental causation by ignoring or misconstruing the role of causal powers in psychological taxonomy. Connecting this up with the methodological formulations that have had influence in cognitive science itself, individualism has been claimed to be a minimal constraint on arriving at psychological explanations that locate the mind suitably in the physical world, a psychology that taxonomizes its entities by their causal powers. (We have seen, however, that individualists themselves disagree about what this implies about psychology.)

Functionalism is the view that psychological states and processes should be individuated by their causal or functional roles – that is, by their place within the overall causal economy of the organism – and it has been common to suppose that these functional or causal roles are individualistic. Certainly, these causal roles can be understood in different ways, but the two (complementary) ways most prevalent in cognitive science – in terms of the notion of computation (e.g., Fodor 1980; Pylyshyn 1984), and in terms of the idea of analytical decomposition (e.g., Dennett 1978; Cummins 1983) – lend themselves to an individualistic reading. Computational processes, operating solely on the *syntactic* properties of mental states, have been plausibly thought to be individualistic; and it is natural to think of analytical decomposition as beginning with a psychological capacity (e.g., memory, depth perception, reasoning) and seeking the intrinsic properties of the organism in virtue of which it instantiates that capacity.

Despite their prima facie plausibility, however, neither of these connections – between physicalism and individualism, and between functionalism and individualism – is unproblematic, and in fact I think that upon closer examination neither purported inference holds. These claims can be explored more fully by examining explicit arguments for individualism that specify these connections more precisely.

11.5 The Appeal to Causal Powers

An argument for individualism that has been widely discussed derives from chapter 2 of Fodor's *Psychosemantics* (1987). Although a series of related criticisms (van Gulick 1989; Egan 1991; Wilson 1992, 1995: ch. 2) seem to me decisive in showing the argument to be fatally flawed, the argument itself taps into an intuition, or perhaps a cluster of intuitions, running deep in the philosophical community. The argument itself is easy to state. Taxonomy or individuation in the sciences in general satisfies a generalized version of individualism: sciences taxonomize the entities they posit and discover by their causal powers. Psychology and the cognitive sciences should be no exception here. But the causal powers of

anything supervene on that thing's intrinsic, physical properties. Thus, scientific taxonomy, and so psychological taxonomy, must be individualistic.

One way to elicit the problem with this argument is to ask what it is that makes the first premise (about scientific taxonomy in general) true. Given the naturalistic turn supposedly embraced by those working in contemporary philosophical psychology, one would think that the support here comes from an examination of actual taxonomic practice across the sciences. However, once one does turn to look at these practices, it is easy to find a variety of sciences that *don't* taxonomize "by causal powers;" rather, they individuate their kinds *relationally*, where often enough it is the *actual* relations that determine kind membership. Examples often cited here include species in evolutionary biology, which are individuated phylogenetically (and so historically), and continents in geology, whose causal powers are pretty much irrelevant to their identity as continents (see Burge 1986a). The problem is particularly acute in the context of this argument for individualism, since a further premise in the argument states that a thing's causal powers supervene on that thing's intrinsic properties, and so one cannot simply stipulate that individuation in these sciences is "by causal powers" in some extended sense of "causal powers." (If one does that, then "causal powers" no longer so supervene.)

The intuition that persists despite an acknowledgment that the argument itself is flawed in something like the way identified above is that individualism *does* articulate a constraint for the explanation of cognition that sciences more generally satisfy, one that would make for a physicalistically respectable psychology (e.g., see Walsh 1999). My view is that this intuition itself seriously underestimates the diversity in taxonomic and explanatory practice across the sciences (see Wilson 2000b), and it simply needs to be given up. Attempts to revitalize this sort of argument for individualism proceed by making the sorts of a priori assumptions about the nature of scientific taxonomies and explanations that are reminiscent of the generalized, rational reconstructions of scientific practice that governed logical positivist views of science, and this should sound alarm bells for any self-professed contemporary naturalistic philosopher of mind.

11.6 Externalism and Metaphysics

What, then, of the more general, putative connection between physicalism and individualism? If the denial of individualism could be shown to entail the denial of a plausibly general version of physicalism, then I think that externalism would itself be in real trouble. But like the individualist's appeal to causal powers and scientific taxonomy, I suspect that the move from the general intuitions that motivate such an argument to the argument itself will itself prove problematic. For example, externalists can respect the physicalist slogan "no psychological difference without a physical difference" because the relevant physical differences lie beyond the boundary of the individual; attempts to refine this slogan (e.g., no

psychological difference without a here-and-now physical difference) are likely either to beg the question against the externalist or invoke a construal of physicalism that is at least as controversial as individualism itself.

What is true is that externalists themselves have not been as attentive to the metaphysical notions at the core of contemporary materialism as they could have been, and when they have so attended they have sometimes sounded opposed to physicalism. The most prominent case here is Burge's (1979) original discussion of the implications of individualism for related views about the mind, where he claimed that the rejection of individualism implied the rejection of both type-type and token-token identity theories of the mind, these being two of the major forms of materialism.

To my mind, the most under-discussed of these notions is that of *realization*. Although it has been common to express materialism as entailing that all mental states are realized as physical states, and to take the relevant physical states to be states of the brain, there has been little general discussion of the properties of this relation of realization, or of the properties of realizer states (see Shoemaker 2000, Gillett 2002, though). This creates a problem for externalists, since the standard view of realization smuggles in an individualistic bias. On this standard view, realizers are held to be both *metaphysically sufficient* for the states they realize and *physically constitutive* of the individuals with the realized properties. Denying the second of these conjuncts, as I think an externalist should, creates space for the idea that mental states have a *wide realization*, an option that I have attempted elsewhere to defend in the context of a more general discussion of realization (Wilson 2001).

11.7 The Debate Over Marr's Theory of Vision

I have already said that individualism receives prima facie support from the computational and representational theories of mind, and thus from the cognitive science community in which those theories have been influential. But I have also indicated that I think that the claim that a truly explanatory cognitive science will be individualistic has an epistemic basis more like a gesture than a proof. One way to substantiate this second view in light of the first is to turn to examine the continuing philosophical debate over whether David Marr's celebrated theory of early vision is individualistic. Apart from the intrinsic interest of the debate itself, our examination here will also help to elicit some of the broader issues about the mind to which the individualism issue is central, including the nature of computation and representation.

In the final section of "Individualism and the Mental," Burge had suggested that his thought experiment and the conclusion derived from it – that mental content and thus mental states with content were not individualistic – had implications for computational explanations of cognition. These implications were

twofold. First, purely computational accounts of the mind, construed individualistically, were inadequate; and second, insofar as such explanations did appeal to a notion of mental content, they would fail to be individualistic. It is the latter of these ideas that Burge pursued in "Individualism and Psychology" (1986a), in which he argued, strikingly, that Marr's theory of vision was *not* individualistic. This was the first attempt to explore a widely respected view within cognitive science vis-à-vis the individualism issue, and it was a crucial turning point in moving beyond the cognitive science gesture toward a style of argument that really does utilize empirical practice in cognitive science itself.

As has often been pointed out, what is called "Marr's theory of vision" is an account of a range of processes in early or "low-level" vision that was developed by Marr and colleagues, such as Ellen Hildreth and Tomas Poggio, at the Massachussetts Institute of Technology from the mid-1970s. These processes include stereopsis, the perception of motion, and shape and surface perception, and the approach is explicitly computational. Marr's *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (1982), published posthumously after Marr's tragic early death in 1980, became the paradigm expression of the approach, particularly for philosophers, something facilitated by Marr's comfortable blend of computational detail with broad-brushed, programmatic statements of the perspective and implications of his approach to understanding vision. Since the publication of Marr's book, work on his theory of vision has continued, being extended to cover the processes constituting low-level vision more extensively (e.g., see Hildreth and Ullman 1989). Interestingly, by and large, the philosophical literature on individualism that appeals to Marr's theory has been content to rely almost exclusively on his *Vision* in interpreting the theory.

Critical to the computational theory that Marr advocates is a recognition of the different levels at which one can – indeed, for Marr, must – study vision. According to Marr, there are three levels of analysis to pursue in studying an information-processing device. First, there is the level of the computational theory (hereafter, the *computational level*), which specifies the goal of the computation, and at which the device itself is characterized in abstract, formal terms as "mapping from one kind of information to another" (1982: 24). Second is the level of representation and algorithm (hereafter, the *algorithmic level*), which selects a "representation for the input and output and the algorithm to be used to transform one into the other" (ibid.: 24–5). And third is the level of hardware implementation (hereafter, the *implementational level*), which tells us how the representation and algorithm are realized physically in the actual device.

Philosophical discussions, like Marr's own discussions, have been focused on the computational and algorithmic levels for vision, what Marr himself (ibid.: 23) characterizes, respectively, as the "what and why" and "how" questions about vision. As we will see, there is particular controversy over what the computational level involves. In addition to the often-invoked trichotomy of levels at which an informational-processing analysis proceeds, there are two further interesting dimensions to Marr's approach to vision that have been somewhat neglected in

the philosophical literature. These add some complexity not only to Marr's theory, but also to the issue of how "computation" and "representation" are to be understood in it.

The first is the idea that visual computations are performed sequentially in *stages of computational inference*. Marr states that the overall goal of the theory of vision is "to understand how descriptions of the world may efficiently and reliably be obtained from images of it" (ibid.: 99). Marr views the inferences from intensity changes in the retinal image to full-blown three-dimensional descriptions as proceeding via the construction of a series of preliminary representations: the raw primal sketch, the full primal sketch, and the $2\frac{1}{2}$-D sketch. Call this the *temporal dimension* to visual computation. The second is that visual processing is subject to modular design, and so particular aspects of the construction of 3-D images – stereopsis, depth, motion, etc. – can be investigated in principle independently. Call this the *modular dimension* to visual computation.

A recognition of the temporal and modular dimensions to visual computation complicates any discussion of what "the" computational and algorithmic levels for "the" process of vision are. Minimally, in identifying each of Marr's three levels, we need first to fix at least the modular dimension to vision in order to analyze a given visual *process*; and to fix at least the temporal dimension in order to analyze a given visual *computation*.

Burge's argument that Marr's theory is not individualistic is explicitly and fully presented in the following passage:

> (1) The theory is intentional. (2) The intentional primitives of the theory and the information they carry are individuated by reference to contingently existing physical items or conditions by which they are normally caused and to which they normally apply. (3) So if these physical conditions and, possibly, attendant physical laws were regularly different, the information conveyed to the subject and the intentional content of his or her visual representations would be different. (4) It is not incoherent to conceive of relevantly different (say, optical) laws regularly causing the same non-intentionally, individualistically individuated physical regularities in the subject's eyes and nervous system. . . . (5) In such a case (by (3)) the individual's visual representations would carry different information and have different representational content, though the person's whole non-intentional physical history . . . might remain the same. (6) Assuming that some perceptual states are identified in the theory in terms of their informational or intentional content, it follows that individualism is not true for the theory of vision. (1986a: 34)

The second and third premise make specific claims about Marr's theory of vision, while the first premise, together with (4) and (5), indicate the affinity between this argument and the Twin Earth-styled argument of Burge's that we discussed earlier.

Burge himself concentrates on defending (2)–(4), largely by an appeal to the ways in which Marr appears to rely on "the structure of the real world" in articulating both the computational and algorithmic levels for vision. Marr certainly does make a number of appeals to this structure throughout *Vision*. For example, he says

> The purpose of these representations is to provide useful descriptions of aspects of the real world. The structure of the real world therefore plays an important role in determining both the nature of the representations that are used and the nature of the processes that derive and maintain them. An important part of the theoretical analysis is to make explicit the physical constraints and assumptions that have been used in the design of the representations and processes. (1982: 43; cf. also pp. 68, 103–5, 265–6)

And Marr does claim that the representational primitives in early vision – such as "blobs, lines, edges, groups, and so forth" – that he posits "correspond to real physical changes on the viewed surface" (ibid.: 44). Together, these sorts of comment have been taken to support (2) and (3) in particular.

Much of the controversy over how to interpret Marr's theory turns on whether this is the correct way to understand his appeals to the "structure of the real world." There are at least two general alternatives to viewing such comments as claiming the importance of the beyond-the-head world for the computational taxonomy of visual states.

The first is to see them as giving the real world a role to play *only* in constructing what Marr calls the computational theory. Since vision is a process for extracting information from the world in order to allow the organism to act effectively in that world, clearly we need to know something of the structure of the world in our account of what vision is *for*, what it is that vision *does*, what function vision is *designed to perform*. If this is correct, then it seems possible to argue that one does *not* need to look beyond the head in constructing the theory of the representation and algorithm. As it is at this level that visual states are taxonomized qua the objects of computational mechanisms, Marr's references to the "real world" do not commit him to an externalist view of the taxonomy of visual states and processes.

The second is to take these comments to suggest merely a *heuristic* role for the structure of the real world, not only in developing a computational taxonomy but in the computational theory of vision more generally. That is, turning to the beyond-the-head world is merely a useful short-cut for understanding how vision works and the nature of visual states and computations, either by providing important *background* information that allows us to understand the representational primitives and thus the earliest stages of the visual computation, or by serving as interpretative lenses that allow us to construct a *model* of computational processes in terms that are meaningful. Again, as with the previous option, the beyond-the-head world plays only a peripheral role within computational vision, even if Marr at times refers to it prominently in outlining his theory.

Individualists have objected to Burge's argument in two principal ways. First, Segal (1989) and Matthews (1988) have both in effect denied (2), with Segal arguing that these intentional primitives (such as edges and generalized cones) are better interpreted within the context of Marr's theory as individuated by their *narrow content*. Second, Egan (1991, 1992, 1995, 1999) has more strikingly

denied (1), arguing that, qua computational theory, Marr's theory is not intentional at all. Both objections are worth exploring in detail, particularly insofar as they highlight issues that remain contentious in contemporary discussions. In fact, the discussion of Marr's theory raises more foundational questions than it solves about the nature of the mind and how we should investigate it.

Segal points out that there are two general interpretations available when one seeks to ascribe intentional contents to the visual states of two individuals. First, one could follow Burge and interpret the content of a given visual state in terms of what normally causes it. Thus, if it is a crack in a surface that plays this role, then the content of the corresponding visual state is "crack;" if it is a shadow in the environment that does so, then the content of the visual state is "shadow." This could be so even in the case of doppelgangers, and so the visual states so individuated are not individualistic. But second and alternatively, one could offer a more liberal interpretation of the content of the visual states in the two cases, one that was neutral as to the cause of the state, and to which we might give the name "crackdow" to indicate this neutrality. This content would be shared by doppelgangers, and so *would* be individualistic.

The crucial part of Segal's argument is his case for preferring the second of these interpretations, and it is here that one would expect to find an appeal to the specifics of Marr's theory of vision. While some of Segal's arguments here do so appeal, he also introduces a number of quite general considerations that have little to do with Marr's theory in particular. For example, he points to the second interpretation as having "economy on its side" (1989: 206), thus appealing to considerations of simplicity, and says:

> The best theoretical description will *always* be one in which the representations fail to specify their extensions at a level that distinguishes the two sorts of distal cause. It will *always* be better to suppose that the extension includes both sorts of thing. (ibid.: 207; my emphasis)

Why "always"? Segal talks generally of the "basic canons of good explanation" (ibid.) in support of his case against externalism, but as with the appeals to the nature of scientific explanation that turned on the idea that scientific taxonomy and thus explanation individuates by "causal powers," here we should be suspicious of the level of generality (and corresponding lack of substantive detail) at which scientific practice is depicted. Like Burge's own appeal to the objectivity of perceptual representation in formulating a general argument for externalism (1986a: section 3; 1986b), these sorts of a priori appeals seem to me to represent gestural lapses entwined with the more interesting, substantive, empirical arguments over individualism in psychology.

When Segal does draw more explicitly on features of Marr's theory, he extracts three general points that are relevant for his argument that the theory is individualistic: each attribution of a representation requires a "bottom-up account" (1989: 194), a "top-down motivation" (ibid.: 195) and is "checked against behavioral

evidence" (ibid.: 197). Together, these three points imply that positing representations in Marr's theory does not come cheaply, and indeed is tightly constrained by overall task demands and methods. The first suggests that any higher-level representations posited by the theory must be derived from lower-level input representations; the second that all posited representations derive their motivation from their role in the overall perceptual process; and the third that "intentional contents are inferred from discriminative behavior" (ibid.: 197).

Segal uses the first assumption to argue that since the content of the earliest representations – "up to and including zero-crossings" (ibid.: 199) – in doppelgangers are the same, there is a prima facie case that downstream, higher-level representations must be the same, unless a top-down motivation can be given for positing a difference. But since we are considering doppelgangers, there is no behavioral evidence that could be used to diagnose a representational difference between the two (Segal's third point), and so no top-down motivation available. As he says, "[t]here would just be no theoretical point in invoking the two contents [of the twins], where one would do. For there would be no theoretical purpose served by distinguishing between the contents" (ibid.: 206).

How might an externalist resist this challenging argument? Three different tacks suggest themselves, each of which grants less to Segal than that which precedes it.

First, one could grant the three points that Segal extracts from his reading of Marr, together with his claim that the lowest levels of representation are individualistic, but question the significance of this. Here one could agree that the gray arrays with which Marr's theory begins do, in a sense, represent light intensity values, and that zero-crossings do, in that same sense, represent a sudden change in the light intensity. But these are both merely representations of some state of the retina, not of the world, and it should be no surprise that such intra-organismic representations have narrow content. Moreover, the depth of the intentionality or "aboutness" of such representations might be called into question precisely because they don't involve any causal relation that extends beyond the head; they might be thought to be representational in much the way that my growling stomach represents my current state of hunger. However, once we move to downstream processes, processes that are later on in the temporal dimension to visual processing, *genuinely* robust representational primitives come into play, primitives such as "edge" and "generalized cone." And the contents of states deploying *these* primitives, one might claim, as representations of a state of the world, metaphysically depend on what they correspond to in the world, and so are not individualistic. The plausibility of this response to Segal turns on both the strength of the distinction between a weaker and a stronger sense of "representation" in Marr's theory, and the claim that we need the stronger sense to have states that are representational in some philosophically interesting sense.

Secondly, and more radically, one could allow that all of the representational primitives posited in the theory represent in the same sense, but challenge the claim that the content of *any* of the corresponding states is narrow: it is wide

content all the way out, if you like. The idea that the representational content of states deploying gray arrays and zero-crossings is in fact wide might itself take its cue from Segal's second point – that representations require a top-down motivation – for it is by reflecting on the point of the overall process of constructing reliable, three-dimensional images of a three-dimensional visual world that we can see that even early retinal representations must be representations of states and conditions in the world. This view would of necessity go beyond Marr's theory itself, which is explicitly concerned only with the computational problem of how we infer three-dimensional images from impoverished retinal information, but would be, I think, very much in the spirit of what we can think of as a Gibsonian aspect to Marr's theory (cf. Shapiro 1993).

Thirdly, and least compromisingly, one could reject one or more of Segal's three points about Marr's theory or, rather, the significance that Segal attaches to these points. Temporally later representations *are* derived from earlier representations, but this itself doesn't tell us anything about how to individuate the contents of *either*. Likewise, that Marr himself begins with low-level representations of the retinal image tells us little about whether such representations are narrow or wide. Top-down motivations *are* needed to justify the postulation of representations, but since there is a range of motivations within Marr's theory concerning the overall point of the process of three-dimensional vision, this also gives us little guidance about whether the content of such representations is narrow or wide. Behavioral evidence does play a role in diagnosing the content of particular representations, but since Marr is not a behaviorist, behavioral discrimination does not provide a litmus test for representational difference (Shapiro 1993: 498–503).

This third response seems the most plausible to develop in detail, but it also seems to me the one that implies that there is likely to be no definitive answer to the question of whether Marr's theory employs either a narrow or a wide notion of content, or both or neither. Although Marr was not concerned at all himself with the issue of the intentional nature of the primitives of this theory, the depth of his methodological comments and asides has left us with an embarrassment of riches when it comes to possible interpretations of his theory. This is not simply an indeterminacy about what Marr meant or intended, but one within the computational approach to vision itself, and, I think, within computational psychology more generally. With that in mind, I shall turn now to Egan's claim that the theory is *not intentional at all*, a minority view of Marr's theory that has not, I believe, received its due (cf. critiques of Egan by Butler 1996 and 1998 and Shapiro 1997; see also Chomsky 1995: 55, fn. 25).

At the heart of Egan's view of Marr is a particular view of the nature of Marr's *computational* level of description. Commentators on Marr have almost universally taken this to correspond to what others have called the "knowledge level" (Newell 1980) or the "semantic level" (Pylyshyn 1984) of description, i.e., as offering an *intentional* characterization of the computational mechanisms governing vision and other cognitive processes. Rather than ignoring Marr's computational level, as some (e.g., Shapiro 1997) have claimed she does (supposedly in

order to focus exclusively on Marr's algorithmic level of description), Egan rejects this dominant understanding of the computational level, arguing instead that what makes it a computational level is that it specifies the function to be computed by a given algorithm in precise, mathematical terms. That is, while this level of description is functional, what makes it the first stage in constructing a *computational* theory is that it offers a function-theoretic characterization of the computation, and thus abstracts away from all other functional characterizations. Thus, while vision might have all sorts of functions that can be specified in language relatively close to that of common sense (e.g., it's for extracting information from the world, for perceiving an objective world, for guiding behavior), none of these, in Egan's view, forms a part of Marr's computational level of description. Given this view, the case for Marr's theory being individualistic because computational follows readily:

> A computational theory prescinds from the actual environment because it aims to provide an abstract, and hence completely general, description of a mechanism that affords a basis for predicting and explaining its behavior in any environment, even in environments where what the device is doing cannot comfortably be described as *cognition*. When the computational characterization is accompanied by an appropriate intentional interpretation, we can see how a mechanism that computes a particular mathematical function can, in a particular context, subserve a cognitive function such as vision. (1995: 191).

According to Egan, while an intentional interpretation links the computational theory to our common-sense-based understanding of cognitive functions, it forms no part of the computational theory itself. Egan's view naturally raises questions not only about what Marr meant by the computational level of description but, more generally, about the nature of *computational* approaches to cognition.

There are certainly places in which Marr does talk of the computational level as simply being a high-level functional characterization of what vision is for, and thus primarily as orienting the researcher to pose certain general questions. For example, one of his tables offers the following summary questions that the theory answers at the computational level: "What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?" (1982: 25, fig. 1–4). Those defending the claim that Marr's theory is externalist have typically rested with this broad and somewhat loose understanding of the computational level of the theory (see, e.g., Burge 1986a: 28; Shapiro 1993: 499–500; 1997: 134).

The problem with this broad understanding of the computational level, and thus of computational approaches to cognition, is that while it builds a bridge between computational psychology and more folksy ways of thinking about cognition, it creates a gap within the computational approach between the computational and algorithmic levels. For example, if we suppose that the computational level specifies simply that some visual states have the function of representing

edges, others the function of representing shapes, etc., there is nothing about such descriptions that guides us in constructing *algorithms* that generate the state-to-state transitions at the heart of computational approaches to vision. More informal elaborations of what vision is for, or of what it evolved to do, do little by themselves to bridge this gap.

The point here is that computational specifications themselves are a very special kind of functional characterization, at least when they are to be completed or implemented in automatic, algorithmic processes. Minimally, proponents of the broad interpretation of computational approaches to cognition need either to construe the computational level as encompassing but going beyond the function-theoretic characterizations of cognitive capacities that Egan identifies, or they must allocate those characterizations to the algorithmic level. The latter option simply exacerbates the "gap" problem identified above. But the former option seems to me to lump together a variety of quite different things under the heading of "the computational level," and subsequently fails to recognize the constraints that computational assumptions bring in their wake. The temporal and modularity dimensions to Marr's theory exacerbate the problem here.

There is a large issue lurking here concerning how functionalism should be understood within computational approaches to cognition, and correspondingly how encompassing such approaches really are. Functionalism has usually been understood as offering a way to reconcile our folk psychology, our manifest image (Sellars 1962) of the mind, with the developing sciences of the mind, even if that reconciliation involves revising folk psychology along individualistic lines (e.g., factoring it into a *narrow* folk psychology via the notion of narrow content). And computationalism has been taken to be one way of specifying what the relevant functional roles are: they are "computational roles." But if Egan is right about Marr's understanding of the notion of computation as a function-theoretic notion, and we accept the view that this understanding is shared in computational approaches to cognition more generally, then the corresponding version of functionalism about the mind must be correspondingly function-theoretic: it must not only "prescind from the actual environment," as she claims the computational level must do, but also from the sort of *internal* causal role that functionalists have often appealed to. Cognitive mechanisms, on this view, take mathematically characterizable inputs to deliver mathematically characterizable outputs, and qua computational devices, that is all. Any prospects for the consilience of our "two images" must lie elsewhere.

In arguing for the non-intentional character of Marr's theory of vision, Egan presents an austere picture of the heart of computational psychology, one that accords with the individualistic orientation of computational cognitive science as it has traditionally been developed (cf. Chomsky 1995), even if computational psychologists have sometimes (e.g., Pylyshyn 1984) attempted to place their theories within more encompassing contexts. One problem with such a view of computation, as Shapiro (1997: 149) points out, is that a computational theory of X tells us very little about the nature of X, including information sufficient to

individuate X as (say) a *visual* process at all. While Egan (1999) seems willing to accept this conclusion, placing this sort of concern outside of computational theory proper, this response highlights a gap between computational theory, austerely construed, and the myriad of theories – representational, functional, or ecological in nature – with which such a theory must be integrated for it to constitute a complete, mechanistic account of any given cognitive process. The more austere the account of computation, the larger this gap becomes, and the less a computational theory contributes to our understanding of cognition. One might well think that Egan's view of computational theory in psychology errs on the side of being *too* austere in this respect.

11.8. Exploitative Representation and Wide Computationalism

As a beginning on an alternative way of thinking about computation and representation, consider an interesting difference between individualistic and externalist interpretations of Marr's theory that concerns what it is that Marrian computational systems have built into them. Individualists about computation, such as Egan and Segal, hold that they incorporate various *innate assumptions* about what the world is like. This is because the process of vision involves recovering 3-D information from a 2-D retinal image, a process that without further input would be underdetermined. The only way to solve this underdetermination problem is to make innate assumptions about the world. The best known of these is Ullman's *rigidity assumption*, which says that "any set of elements undergoing a two-dimensional transformation has a unique interpretation as a rigid body moving in space and hence should be interpreted as such a body in motion" (1979: 146). The claim that individualists make is that assumptions like this are part of the computational systems that drive cognitive processing. This is the standard way to understand Marr's approach to vision.

Externalists like Shapiro have construed this matter differently. Although certain assumptions must be true of the world in order for our computational mechanisms to solve the underdetermination problem, these are simply assumptions that are *exploited* (Shapiro 1997: 135, 143; cf. Rowlands 1999) by our computational mechanisms, rather than innate in our cognitive architecture. That is, the assumptions concern the relationships between features of the external world, or between properties of the internal, visual array and properties of the external world, but those assumptions are not themselves encoded in the organism. To bring out the contrast between these two views, consider a few simple examples.

An odometer keeps track of how many miles a car has traveled, and it does so by counting the number of wheel rotations and being built so as to display a number proportional to this number. One way in which it could do this would be for the assumption that 1 rotation = x meters to be part of its calculational machinery; another way of achieving the end would be for it to be built so as

simply to record x meters for every rotation, thus exploiting the fact that 1 rotation = x meters. In the first case it encodes a representational assumption, and uses this to compute its output; in the second, it contains no such encoding but instead *uses* an existing relationship between its structure and the structure of the world. In either case, if it finds itself in an environment in which the relationship between rotations to distance traveled is adjusted (e.g., larger wheels, or being driven on a treadmill), it will not function as it is supposed to, and will misrepresent the distance traveled.

Consider two different strategies for learning how to hit a baseball that is falling vertically to the ground. Since the ball accelerates at 9.8 ms^{-2}, there is a time lag between swinging and hitting. One could either assume that the ball is falling (say, at a specific rate of acceleration), and then use this assumption to calculate when one should swing; alternatively, one could simply aim a certain distance below where one perceives the ball at the time of swinging (say, two feet). In this latter case one would be exploiting the relationship between acceleration, time, and distance without having to encode that relationship in the assumptions one brings to bear on the task.

The fact that there are these two different strategies for accomplishing the same end should, minimally, make us wary of accepting the claim that innate assumptions are the *only* way that a computational system could solve the underdetermination problem. But I also want to develop the idea that our perceptual system in particular and our cognitive systems more generally typically exploit rather than encode information about the world and our relationship to it, as well as say something about where Marr himself seems to stand on this issue (see also Wilson, forthcoming).

An assumption that Egan makes and that is widely shared in the philosophical literatures both on individualism and computation is that at least the *algorithmic* level of description within computational psychology is individualistic. The idea here has, I think, seemed so obvious that it has seldom been spelled out: algorithms operate on the syntactic or formal properties of symbols, and these are intrinsic to the organisms instantiating the symbols. We might challenge this neither by disputing how much is built into Marr's computational level, nor by squabbling over the line between Marr's computational and algorithmic levels, but, rather, by arguing that computations themselves can extend beyond the head of the organism and involve the relations between individuals and their environments. This position, which holds that at least some of the computational systems that drive cognition, especially human cognition, reach beyond the limits of the organismic boundary, I have elsewhere (1994; 1995: ch. 3) called *wide computationalism*, and its application to Marr's theory of vision marks a departure from the parameters governing the standard individualist-externalist debate over that theory. Wide computationalism constitutes one way of thinking about the way in which cognition, even considered computationally, is "embedded" or "situated" in its nature (cf. also Hutchins 1995; McClamrock 1995), and it provides a framework within which an exploitative conception of representation can be pursued.

The basic idea of wide computationalism is simple. Traditionally, the sorts of computation that govern cognition have been thought to begin and end at the skull. But why think that the skull constitutes a magic boundary beyond which true computation ends and mere causation begins? Given that we are creatures embedded in informationally rich and complex environments, the computations that occur inside the head are an important part but are not exhaustive of the corresponding computational systems. This perspective opens up the possibility of exploring computational units that include the brain as well as aspects of the brain's beyond-the-head environment. Wide computational systems thus involve minds that literally extend beyond the confines of the skull into the world.

One way to bring out the nature of the departure made by wide computationalism within the individualism debate draws on a distinction between a *locational* and a *taxonomic* conception of psychological states (see also Wilson 2000a; cf. Rowlands 1999: chs. 2–3). Individualists and externalists are usually presented as disagreeing over how to taxonomize or individuate psychological states, but both typically (though not always) presume that the relevant states are what we might call *locationally individualistic*: they are located within the organismic envelope. What individualists and externalists typically disagree about is whether in addition to being locationally individualistic, psychological states must also be taxonomically individualistic. Wide computationalism, however, rejects this assumption of locational individualism by claiming that some of the "relevant states" – some of those that constitute the relevant computational system – are located not in the individual's head but in her environment.

The intuitive idea behind wide computationalism is easy enough to grasp, but there are two controversial claims central to defending wide computationalism as a viable model for thinking about and studying cognitive processing. The first is that it is sometimes appropriate to offer a formal or computational characterization of an organism's environment, and to view parts of the brain of the organism, computationally characterized, together with this environment so characterized, as constituting a unified computational system. Without this being true, it is difficult to see wide computationalism as a coherent view. The second is that this resulting mind–world computational system itself, and not just the part of it inside the head, is genuinely cognitive. Without this second claim, wide computationalism would at best present a zany way of carving up the computational world, one without obvious implications for how we should think about real cognition in real heads. Rather than attempting to respond to each of these problems in the space available, I shall turn to the issue of how this general perspective on representation and computation sits with Marr's theory of vision.

As we have seen, Marr himself construes the task of a theory of vision to be to show how we extract visual information from "arrays of image intensity values as detected by the photoreceptors in the retina" (1982: 31). Thus, as we have already noted, the problem of vision *begins* with retinal images, not with properties of the world beyond those images, and "the true heart of visual perception is the inference *from* the structure of an image about the structure of the real world

outside" (ibid.: 68; my emphasis). Marr goes on to characterize a range of physical constraints that hold true of the world that "make this inference possible" (ibid.), but he makes it clear that "the constraints are used by turning them into an assumption that may or may not be internally verifiable" (ibid.: 104). For all of Marr's talk of the importance of facts about the beyond-the-head world for constructing the computational level in a theory of vision, this is representative of how he conceives of that relevance (e.g., ibid.: 43, 68, 99, 103–5, 265–6). It seems to me clear that, in terms that I introduced earlier in this section, Marr himself adopts an encoding view of computation and representation, rather than an exploitative view of the two. The visual system is, according to Marr, a locationally individualistic system.

Whatever Marr's own views here, the obvious way to defend a wide computational interpretation of his theory is to resist his inference from "*x* is a physical constraint holding in the world" to "*x* is an assumption that is encoded in the brain." This is, in essence, what I have previously proposed one should do in the case of the multiple spatial channels theory of form perception pioneered by Campbell and Robson (1968). Like Marr's theory of vision, which in part builds on this work (see esp. Marr 1982: 61–4), this theory has usually been understood as postulating a locationally individualistic computational system, one that begins with channels early in the visual pathway that are differentially sensitive to four parameters: orientation, spatial frequency, contrast, and spatial phase. My suggestion (Wilson 1994; 1995: ch. 3) was to take seriously the claim that any visual scene (in the world) can be decomposed into these four properties, and so see the computational system itself as extending into the world, with the causal relationship between stimulus and visual channels itself modeled by transition rules. Rather than simply having these properties encoded in distinct visual channels in the nervous system, view the in-the-head part of the form perception system as exploiting formal properties in the world beyond the head. With respect to Marr's theory, there is a respect in which this wide computational interpretation is *easy* to defend, and another in which it is *difficult* to defend.

The first of these is that Marr's "assumptions," such as the spatial coincidence assumption (1982: 70) and the "fundamental assumption of stereopsis" (ibid.: 114), typically begin as physical constraints that reflect the structure of the world; in the above examples, they begin as the constraint of spatial localization (ibid.: 68–9) and three matching constraints (ibid.: 112–14). Thus, the strategy is to argue that the constraints *themselves*, rather than their derivative encoding, play a role in defining the computational system, rather than simply filling a heuristic role in allowing us to offer a computational characterization of a locationally individualistic cognitive system.

The corresponding respect in which a wide computational interpretation of Marr's theory is difficult to defend is that these constraints themselves do not specify what the computational primitives are. One possibility would simply be to attribute the primitives that Marr ascribes to the *image* to features of the scenes perceived themselves, but this would be too quick. For example, Marr considers

zero-crossings to be steps in a computation that represent sharp changes in intensity in the image, and while we could take them to represent intensity changes in the stimuli in the world, zero-crossings themselves are located somewhere early in the in-the-head part of the visual system, probably close to the retina. A better strategy, I think, would be to deflate the interpretation of the retinal image and look "upstream" from it to identify richer external structures in the world, structures which satisfy the physical constraints that Marr postulates. That is, one should extend the temporal dimension to Marr's theory so that the earliest stages in basic visual processes *begin in the world*, not in the head. Since the study of vision has been largely conducted within an overarching individualistic framework, this strategy would require recasting the theory of vision itself so that it ranges over a process that causally extends beyond the retinal image (see also Rowlands 1999: ch. 5).

11.9 Narrow Content and Marr's Theory

Consider the very first move in Segal's argument for the conclusion that Marr's theory of vision is individualistic, the claim that there are two general interpretations available when one seeks to ascribe intentional contents to the visual states of two individuals: one "restrictive" (Burge's) and one "liberal" (Segal's). Something like these two general alternatives was implicit in the basic Twin Earth cases with which we – and the debate over individualism – began; the idea that twins must share some intentional state about watery substances (or about arthritis-like diseases, in Burge's standard case) is the basis for attempts to articulate a notion of *narrow* content, i.e., intentional content that does supervene on the intrinsic, physical properties of the individual. I have elsewhere (Wilson 1995: ch. 9) expressed my skepticism about such attempts, and here I want to tie this skepticism to the innocuous-looking first step in Segal's interpretation.

This first step in Segal's interpretation, the presupposition of a liberal interpretation for Marr's theory, and a corresponding view of the original Twin Earth cases in general, are themselves questionable. Note first that the representations that we might, in order to make their disjunctive content perspicuous, label "crackdow" or "water or twater," *do* represent their reliable, environmental causes: "crackdow" is reliably caused by cracks or shadows, and has the content crack or shadow; similarly for "water or twater." But then this disjunctive content is a species of wide, not narrow content, as Egan (1995: 195) has pointed out. In short, although being shared by twins is necessary, it is not sufficient for mental content to be narrow.

To press further, if the content of one's visual state is to be individualistic, it must be shared by doppelgangers *no matter how different their environments.* Thus, the case of "twins" is merely a heuristic for thinking about a potentially infinite number of individuals. But then the focus on a content shared by two

individuals, and thus on a content that is neutral between two environmental causes, represents a misleading simplification insofar as the content needed won't simply be "crackdow," but something more wildly disjunctive, since there is a potentially infinite number of environments that might produce the same intrinsic, physical state of the individual's visual system as (say) cracks do in the actual world (see also Egan 1991: 200, fn. 35). It is not that we can't simply make up a name for the content of such a state (we can: call it "X"), but that it is difficult to view a state so individuated as being *about anything*. And if being about something is at the heart of being intentional, then this calls into question the status of such narrowly individuated states as intentional states.

Segal (1991: 490) has claimed that the narrow content of "crackdow," or by implication "water or twater," need not be disjunctive, just simply more encompassing than, respectively, crack or water (see also Segal 2000). But casting the above points in terms of disjunctive content simply makes vivid the general problems that (1) the individuation of states in terms of their content still proceeds via reference to what does or would cause them to be tokened; and (2) once one prescinds from a conception of the cognitive system as embedded in and interacting with the actual world in thinking about how to taxonomize its states, it becomes difficult to delineate clearly those states as intentional states with some definite content. As it is sometimes put, narrow content becomes *inexpressible*. Two responses might be made to this second objection.

First, one might concede that, strictly speaking, narrow content *is* inexpressible, but then point out ways of "sneaking up on it" (Fodor 1987: 52). One might do so by talking of how one can "anchor" narrow content to wide content (ibid.: 50–3); or of how to specify the *realization conditions* for a proposition (Loar 1988). But these suggestions, despite their currency, seem to me little more than whistling in the dark, and the concession on which they rest, fatal. All of the ways of "sneaking up on" narrow content involve using wide contents in some way. Yet if wide content is such a problematic notion (because it is not individualistic), then surely the problem spreads to any notion, such as snuck-up-on narrow content, for whose intelligibility the notion of wide content is crucial.

Moreover, if narrow content really is inexpressible, then the idea that it is this notion that is central to psychological explanation as it is actually practiced, and this notion that does or will feature in the natural kinds and laws of the cognitive sciences, cannot reasonably be sustained. Except in Douglas Adamsesque spoofs of science, there are no sciences whose central explanatory constructs are inexpressible. Moreover, this view would make the claim that one arrives at the notion of narrow content via an examination of actual explanatory practice in the cognitive sciences extremely implausible, since if narrow content is inexpressible, then one won't be able to find it expressed in any existing psychological theory. In short, and in terms that I introduced earlier, the idea that snuck-up-on narrow content is what cognitive science needs or uses represents a reversion to the cognitive science gesture.

Secondly, it might be claimed that although it is true that it *is* difficult for common-sense folk to come up with labels for intentional contents, those in the

relevant cognitive sciences can and do all the time, and we should defer to them. For example, one might claim that many if not all of the representational primitives in Marr's theory, such as blob, edge, and line, have narrow contents. These concepts, like many scientific terms, are technical and, as such, may bear no obvious relationship to the concepts and terms of common sense, but they still allow us to see how narrow content can be expressed. One might think that this response has the same question-begging feel to it as does the claim that our folk psychological states are themselves narrow. However, the underdetermination of philosophical views by the data of the scientific theories, such as Marr's, that they interpret remains a problem for both individualists and externalists alike here. As my discussion of exploitative representation and wide computation perhaps suggests, my own view is that we need to reinvigorate the ways in which the computational and representational theories of mind have usually been construed within cognitive science. If this can be done in more than a gestural manner, then the issue of the (in)expressibility of narrow content will be largely moot.

11.10 Individualism and the Problem of Self-knowledge

Thus far, I have concentrated on discussions of individualism and externalism in contemporary philosophy of mind with a primary affinity to cognitive science. It is testimony to the centrality of individualism and externalism for philosophy more generally – quite apart from their relevance to empirical cognitive science – that there is a variety of discussions that explore the relationship between these positions and traditional issues in the philosophy of mind and philosophy more generally. The most interesting of these seem to me to cluster around three related epistemological issues: self-knowledge, a priori knowledge, and skepticism.

Basic to self-knowledge is knowledge of one's own mind, and traditionally this knowledge has been thought to involve some form of privileged access to one's own mental states. This first-person privileged access has often been understood in terms of one or more distinctive properties that the resulting second-order mental states have. These states, such as my belief that I believe that the Earth goes around the sun, have been claimed to be *infallible* (i.e., incapable of being false or mistaken), which would imply that simply having the second-order belief guarantees that one has the first-order belief that is its object; or *incorrigible* (i.e., even if mistaken, incapable of being corrected by anyone other than the person who has them), which would at least imply that they have a form of epistemic security that other types of mental state lack. In either case, there is an asymmetry between knowledge of one's own mind and knowledge of the minds of others, as well as knowledge of other things in the world. Indicative of the depth of these asymmetries in modern philosophy is the fact that an introduction to epistemology, particularly one with a historical slant, that reflects on skepticism, will likely introduce the *problem of other minds* and the *problem of our knowledge of the external*

world, but not the corresponding *problem of self-knowledge*. Skepticism about one's own mind has seemed to be precluded by the very nature of self-knowledge.

Although the contrast between first- and third-person knowledge of mental states has softened in recent philosophy of mind, it remains part of our common-sense conception of the mind that the ways in which I know about my own mental life are distinctive from the ways in which I know about that of others (cf. Siewert 1998). Thus, not unreasonably, the idea of first-person epistemic privilege survives. Knowledge about one's *self*, about the condition or state of one's mind or body, often enough seems to be simply a matter of introspection, of inward-directed reflection or attention, rather than requiring the collection of evidence through observation or experiment. I simply *feel* my skin itching, or upon attending notice that my toes are squashed up in my shoes; to find out whether *your* skin is itching or whether your toes are squashed up in your shoes, I observe your body and its behavior (including what you *say*), and then draw an inference from that observation to a conclusion about your bodily state. Self-knowledge is *direct*, while knowledge of others is inferential or mediated in some way, based on observation and other forms of evidence. Since one's own mental states are typically the object of first-person thoughts, we are acquainted with our own minds in a way that we are not acquainted with the minds of others.

Individualistic conceptions of the mind have seemed well-suited to making sense of first-person privileged access and the subsequent asymmetry between self-knowledge and knowledge of the mental states of others. If mental states are individuated in abstraction from the beyond-the-individual environment, then there seems to be no problem in understanding how the process of introspection, turning our mind's eye inwards (to use a common metaphor), reveals the content of those states. To invoke the Cartesian fantasy in a way that brings out the asymmetry between self-knowledge and other forms of knowledge, even if there were an evil demon who deceived me about the existence of an external world – including the existence of other people with mental states like mine – the one thing that I could be sure about would be that I am having experiences with a certain content. As it is sometimes put, even if I could be deceived about whether there is really a tree in front of me and thus about whether I am actually seeing a tree, I cannot be deceived about whether *it seems to me that I am seeing a tree*. Thus individualism seems to facilitate a sort of epistemic security for first-person knowledge of one's own mental states that the corresponding third-person knowledge lacks.

Externalism, by contrast, poses a prima facie problem for even the more modest forms of first-person privileged access, and has even been thought to call into question the possibility of any form of self-knowledge. For externalism claims that what mental states *are* is metaphysically determined, in part, by the nature of the world beyond the boundary of the subject of those states. Thus it would seem that in order to know *what* one is thinking, i.e., to know the content of one's mental states, one would have to know something about the world beyond one's self. But this would be to assimilate our first-person knowledge of our own minds to our knowledge of other things, and so deny any *privileged* access that

self-knowledge might be thought to have. It implies that in order to know my own mind I need to know about perhaps difficult-to-discern facts about the nature of the physical or social world in which I live, and so it also suggests that in a range of ordinary cases where we might unreflectively attribute self-knowledge, I don't actually have self-knowledge at all.

We can express the problem here in another way that abstracts from the differences between both specific accounts of privileged access and specific versions of externalism. Whether it be infallible, incorrigible, self-intimating, introspective, or a priori, knowledge of one's own mental states has a special character. Knowing one's own mental states involves, *inter alia*, knowing their contents. Now, according to externalism, the contents of a subject's mental states are metaphysically determined, in part, by facts about her physical or social environment. Knowledge of these facts, however, does not have this special character. But then how is the special character of self-knowledge compatible with the non-special character of worldly knowledge, given the dependence of the former on the latter (see also Ludlow and Martin 1998: 1)? Others have stated the problem more dramatically. For example, Davidson presents it as "a transposed image of Cartesian skepticism" (1987: 94), according to which "[o]ur beliefs about the external view are . . . directed onto the world, but we don't know what we believe" (ibid.), claiming thus that externalism seems to imply that we don't have self-knowledge at all; Heil points out that "if externalism were true, one could not discover a state's intentional properties merely by inspecting that state" (1988: 137), going on to connect this up with Davidson's focus on a "nastier skeptic, one who questions the presumption that we think what we think we think" (ibid.).

The problem can be schematized as a supposedly inconsistent triad of propositions (cf. also McKinsey 1991, whose triad differs; see below). Let P = the contents of our mental states, E = facts about the environment, and let "by introspection" stand in for the distinctive character of self-knowledge:

1 We know P by introspection. (Self-knowledge)
2 P are metaphysically determined in part by E. (Externalism)
3 E are not known by introspection. (Common Sense)

The claim is that one of these three propositions must be given up. If we reject Self-knowledge, then we give up on the idea that we have privileged access to our own minds; if we reject Externalism, then we give up on an independently plausible view of the mind; and if we reject Common Sense, then we make a strange and implausible claim about our knowledge of the physical or social world.

When it is stated so starkly, I think that the right response to the "problem of self-knowledge" is to argue that all three propositions are true, and so consistent, and thus that there is no problem of self-knowledge for an externalist to solve. Their consistency turns on the fact that (1) and (3), which make epistemological claims, are connected only by (2), which makes a metaphysical claim. As a counterexample to the charge of formal inconsistency, consider an instance of the

argument where P = the state of being in pain, and E = a particular, complicated state of the central nervous system. There is no inconsistency in these instances of (1)–(3): we do know that we are in pain by introspection; that state is metaphysically determined by some particular state of our central nervous system; but we don't know about *that* state by introspection. (Or, to put it more carefully: we don't about it qua state of our central nervous system by introspection.) The same is true of our original triad, as well as of variations on those propositions which substitute some other distinctive feature of self-knowledge for "by introspection."

If this is the correct way to represent the supposed problem for externalists, and the basis for an adequate response to that problem, then two features of the problem are worth noting.

The first is that at the heart of the problem is not an externalist view of the mind itself but, rather, any thesis of metaphysical determination, where the determining state is not something that is known in the special way that mental states are known. Since not all of an organism's internal, individualistically individuated states are so known, there is a variation on the problem of self-knowledge that individualists must face, if it is a real problem. Thus, even if one rejects the way of dissolving the problem posed above, a version of the problem of self-knowledge remains for both externalists and individualists to solve. This implies that externalists do *not*, despite initial appearances, face a *special* problem concerning self-knowledge.

The second is that the problem and response so characterized have affinities with a family of problem–response pairs, including on the "problem" side Moore's open question argument and the paradox of analysis, and whose closest relative perhaps is a standard objection to the mind–brain identity theory. Pain, it was claimed, couldn't be identical to C-fiber firing, since one can know that one is in pain but not know that one's C-fibers were firing. And the now-standard response is that such an objection, in attempting to derive an ontological conclusion from epistemological premises, commits a fallacy. Now, as a purportedly inconsistent triad, rather than an argument that draws such a conclusion, the problem of self-knowledge itself does not suffer from this specific problem, although the rejection of externalism as a response to the problem *would* be subject to just this objection. However, the broad affinity here is worth keeping in mind. How adequate one finds the proposed dissolution of the problem of self-knowledge is likely to correlate with how adequate one finds this type of response to this type of objection more generally.

Proponents of the problem of self-knowledge should object to the claim that (1)–(3) adequately expresses the dilemma. In particular, they should (and in fact do) reject (2) as a member of the triad. Rather, the problem of self-knowledge is constituted by the following triad (cf. McKinsey 1991):

A We have a priori knowledge of P. (Self-Knowledge*)
B We have a priori knowledge that P entails E. (Knowledge of Externalism)
C We cannot know E a priori. (Common Sense*)

(A)–(C) *are* inconsistent. But in contrast to (1)–(3), this construal of the problem of self-knowledge can be challenged at *every* point.

First, is an externalist committed to (B)? For an affirmative answer, two prior questions need to be answered affirmatively: according to externalism, (i) do we know that P entails E? and (ii) does P entail E? Take (ii): does externalism claim that, for example, having a mental state with the content "arthritis occurs in the thigh" entail that arthritis does actually occur in the thigh? One reason to think not is that, as we have seen, externalism incorporates the idea that there is a social division of labor in both thought and language, which allows for intentionality even in "vacuous cases": we can think P not because P, but because others think P.

Given, however, that externalism claims that there is a deep, individuative relation between the nature of an individual's mental states and how the world beyond the individual is, some such entailment between P and E seems plausible. This suggests that E needs to be construed in a more nuanced way, encompassing perhaps various disjuncts which together must be true if the externalist's view of the mind is correct. For example, it might be claimed that having the thought that arthritis occurs in the thigh entails either that arthritis does occur in the thigh or that one lives in a linguistic community of a certain character; perhaps more (or more complicated) disjuncts need to be added here (cf. Brown 1995). But then it seems less plausible that "we," i.e., each of us ordinary folk, know (2) so construed, let alone know this a priori. After all, few of us have reflected systematically on what the contents of our thoughts imply about the world; indeed, many of those who have thus reflected – individualists – have concluded that they tell us *nothing* about the character of the world.

This in turn invites the response that to form an inconsistent triad with (A) and (C), (B) need only claim that we *can* have such knowledge, and if externalism is true, and at least some people believe it and what it entails, then that is sufficient to generate the inconsistency.

This seems to me to be a strange way to develop the problem of self-knowledge, since it now sounds like a problem that arises chiefly for the self-knowledge of those versed in the externalism literature, rather than self-knowledge per se. But the real problem here, and the second problem with this construal of the triad, is that the triad now includes a questionable reading of (C). For now (C), even if it *is* a dictate of common sense (and modalized, as (3) is not, this seems doubtful), seems false, since although it is usual for ordinary folk to know about what is mentioned in E through empirical means, and so they don't usually know E a priori, in light of this reading of (B), it seems at least possible that someone could know about E in this fashion. Combined with the reminder that this is not usually how we come to know facts about the empirical world, this concession seems fairly innocuous, and preserves the intuition that self-knowledge is epistemically privileged.

We can see how this construal of the triad undermines its status as a problem for externalism by turning to (A): do we know the contents of our thoughts a

priori? McKinsey conceives of a priori knowledge as knowledge "obtained independently of empirical investigation" (1991: 175), and relies on introspection and reasoning as paradigm processes through which we gain such knowledge. Externalists should be wary of this claim if it is taken to imply that self-knowledge can be gained *completely independently* of empirical investigation of the world; what they can allow, and perhaps all that is needed for (A), is that we know the contents of our mental states on particular occasions without empirically investigating the world on those occasions.

On this reading, (A) is made true by the existence of introspection, while (B)'s truth turns on our ability to follow the arguments for the externalist nature of content and so intentional mental states. While (C) may seem true if we think only of introspection or reasoning alone as means of securing a priori knowledge (in the sense above), it becomes more dubious once we consider introspection and reasoning *together*. Since it is unusual for us to both introspect our own mental states and engage in sophisticated philosophical reasoning using the results of such introspection as premises, the circumstances under which (C) will be falsified are themselves unusual; but (C) nonetheless is, strictly speaking, false.

Note

I should like to thank Gabriel Segal, Frances Egan, and Lawrence Shapiro for reading an earlier version of this review.

References

Block, N. (1986). "Advertisement for a Semantics for Psychology." In P. French, T. Uehling Jr., and H. Wettstein (eds.), *Midwest Studies in Philosophy*, vol. 10 (Philosophy of Mind). Minneapolis: University of Minnesota Press.

Brown, J. (1995). "The Incompatibility of Anti-Individualism and Privileged Access." *Analysis*, 55: 149–56. Reprinted in Ludlow and Martin (1998).

Burge, T. (1979). "Individualism and the Mental." In P. French, T. Uehling Jr., and H. Wettstein (eds.), *Midwest Studies in Philosophy*, vol. 4 (Metaphysics). Minneapolis: University of Minnesota Press.

—— (1986a). "Individualism and Psychology." *Philosophical Review*, 95: 3–45.

—— (1986b). "Cartesian Error and the Objectivity of Perception." In P. Pettit and J. McDowell (eds.), *Subject, Thought, and Context*. Oxford: Oxford University Press. Also in Grimm and Merrill (eds.), *Contents of Thought*. Tucson, AZ: University of Arizona Press (1988).

Butler, K. (1996). "Individualism and Marr's Computational Theory of Vision." *Mind and Language*, 11: 313–37.

—— (1998). "Content, Computation, and Individuation." *Synthese*, 114: 277–92.

Campbell, F. W. and Robson, J. G. (1968). "Application of Fourier Analysis to the Visibility of Gratings." *Journal of Physiology*, 197: 151–66.

Carnap, R. (1928). *The Logical Construction of the World*, trans. by R. George, 1967. Berkeley, CA: University of California Press.

Chomsky, N. (1986). *Knowledge of Language*. New York: Praeger.

—— (1995). "Language and Nature." *Mind*, 104: 1–61.

—— (2000). *New Horizons in the Study of Language and Mind*. New York: Cambridge University Press.

Cosmides, L. and Tooby, J. (1994). "Foreword" to S. Baron-Cohen, *Mindblindness*. Cambridge, MA: MIT Press.

Cummins, R. C. (1983). *The Nature of Psychological Explanation*. Cambridge, MA: MIT Press.

Davidson, D. (1987). "Knowing One's Own Mind." *Proceedings of the American Philosophical Association*. Reprinted in Ludlow and Martin (1998).

Dennett, D. C. (1978). "Artificial Intelligence as Philosophy and as Psychology." In M. Ringle (ed.), *Philosophical Perspectives on Artificial Intelligence*. New York: Humanities Press and Harvester Press. Reprinted in D. C. Dennett, *Brainstorms*. Cambridge, MA: MIT Press.

Egan, F. (1991). "Must Psychology be Individualistic?" *Philosophical Review*, 100: 179–203.

—— (1992). "Individualism, Computation, and Perceptual Content." *Mind*, 101: 443–59.

—— (1995). "Computation and Content." *Philosophical Review*, 104: 181–203.

—— (1999). "In Defense of Narrow Mindedness." *Mind and Language*, 14: 177–94.

Fodor, J. A. (1980). "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology." *Behavioral and Brain Sciences*, 3: 63–73. Reprinted in J. A. Fodor, *Representations*. Sussex: Harvester Press (1981).

—— (1987). *Psychosemantics*. Cambridge, MA: MIT Press.

Gillett, C. (2002). "The Dimensions of Realization: A Critique of the Standard View." *Analysis* (October).

Heil, J. (1988). "Privileged Access." *Mind*, 97: 238–51. Reprinted in Ludlow and Martin (1998).

Hildreth, E. and Ullman, S. (1989). "The Computational Study of Vision." In M. Posner (ed.), *Foundations of Cognitive Science*. Cambridge, MA: MIT Press.

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.

Jackendoff, R. (1991). "The Problem of Reality." Reprinted in his *Languages of the Mind*. Cambridge, MA: MIT Press (1992).

Kripke, S. (1980). *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Loar, B. (1988). "Social Content and Psychological Content." In R. Grimm and D. Merrill (eds.), *Contents of Thought*. Tucson, AZ: University of Arizona Press.

Ludlow, P. and Martin, N. (eds.) (1998). *Externalism and Self-Knowledge*. Palo Alto, CA: CSLI Publications.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco, CA: W. H. Freeman.

Matthews, R. (1988). "Comments on Burge." In R. Grimm and D. Merrill (eds.), *Contents of Thought*. Tucson, AZ: University of Arizona Press.

McClamrock, R. (1995). *Existential Cognition: Computational Minds in the World*. Chicago: University of Chicago Press.

McKinsey, M. (1991). "Anti-Individualism and Privileged Access." *Analysis*, 51: 9–16.

Newell, A. (1980). "Physical Symbol System." *Cognitive Science*, 4: 135–83.
Putnam, H. (1975). "The Meaning of 'Meaning'." In K. Gunderson (ed.), *Language, Mind and Knowledge*. Minneapolis, MN: University of Minnesota Press. Reprinted in H. Putnam, *Mind, Language, and Reality*. New York: Cambridge University Press (1975).
Pylyshyn, Z. (1984). *Computation and Cognition*. Cambridge, MA: MIT Press.
Rowlands, M. (1999). *The Body in Mind*. New York: Cambridge University Press.
Segal, G. (1989). "Seeing What is Not There." *Philosophical Review*, 98: 189–214.
—— (1991). "Defence of a Reasonable Individualism." *Mind*, 100: 485–94.
—— (2000). *A Slim Book About Narrow Content*. Cambridge, MA: MIT Press.
Sellars, W. (1962). "Philosophy and the Scientific Image of Man." In R. Colodny (ed.), *Frontiers of Science and Philosophy*. Pittsburgh: University of Pittsburgh Press. Reprinted in W. Sellars, *Science, Perception, and Reality*. Atascadero, CA: Ridgeview Publishing Company (1963).
Shapiro, L. (1993). "Content, Kinds, and Individualism in Marr's Theory of Vision." *Philosophical Review*, 102: 489–513.
—— (1997). "A Clearer Vision." *Philosophy of Science*, 64: 131–53.
Shoemaker, S. (2000). "Realization and Mental Causation." Reprinted in C. Gillett and B. Loewer (eds.), *Physicalism and it Discontents*. New York: Cambridge University Press.
Siewert, C. (1998). *The Significance of Consciousness*. Princeton, NJ: Princeton University Press.
Stich, S. (1978). "Autonomous Psychology and the Belief-Desire Thesis." *Monist*, 61: 573–91.
—— (1983). *From Folk Psychology to Cognitive Science*. Cambridge, MA: MIT Press.
Ullman, S. (1979). *The Interpretation of Visual Motion*. Cambridge, MA: MIT Press.
Van Gulick, R. (1989). "Metaphysical Arguments for Internalism and Why They Don't Work." In S. Silvers (ed.), *Rerepresentation*. Dordrecht, The Netherlands: Kluwer.
Walsh, D. M. (1999). "Alternative Individualism." *Philosophy of Science*, 66: 628–48.
Wilson, R. A. (1992). "Individualism, Causal Powers, and Explanation." *Philosophical Studies*, 68: 103–39.
—— (1994). "Wide Computationalism." *Mind*, 103: 351–72.
—— (1995). *Cartesian Psychology and Physical Minds: Individualism and the Sciences of the Mind*. New York: Cambridge University Press.
—— (2000a). "The Mind Beyond Itself." In D. Sperber (ed.), *Metarepresentation*. New York: Oxford University Press.
—— (2000b). "Some Problems for 'Alternative Individualism'." *Philosophy of Science*, 67: 671–9.
—— (2001). "Two Views of Realization." *Philosophical Studies*, 104: 1–31.
—— (forthcoming). *The Individual in the Fragile Sciences I: Cognition*.

The Blackwell Guide to Philosophy of Mind

Edited by

Stephen P. Stich and Ted A. Warfield

Blackwell Publishing