Wide Computationalism
Author(s): Robert A. Wilson
Source: *Mind*, New Series, Vol. 103, No. 411 (Jul., 1994), pp. 351-372
Published by: Oxford University Press on behalf of the Mind Association
Stable URL: http://www.jstor.org/stable/2253744
Accessed: 13-03-2016 23:28 UTC

Wide Computationalism

ROBERT A. WILSON

1. Introduction

It has often been thought that individualism in psychology, the view that psychological states must be taxonomized so as to supervene on the intrinsic, physical properties of individuals,[1] receives supports from the computational theory of mind, a view taken by many philosophers and cognitive scientists to be a foundational assumption of contemporary research in cognitive science.[2] The computational theory of mind, or computationalism, can be summarized as the view that psychological processes and mental states are essentially computational. It makes an empirical claim about the nature of cognitive processing and suggests to many a methodological claim about how cognitive psychology, or cognitive science more generally, ought to proceed.

This paper offers a challenge to those who have either argued from computationalism to individualism or thought such an inference plausible by identifying a possibility that has either been overlooked or not treated seriously by proponents of this family of arguments. The possibility is that of *wide computationalism*, and I shall defend both the possibility and the plausibility of wide computationalism in parts of cognitive psychology.

To get a clearer fix both on the type of argument for individualism I have in mind and the nature of the objection I shall pose to it, consider the following explicit argument, which I shall refer to as the *computational argument* for individualism:

[1] Following Stich (1983), I employ a supervenience formulation of individuals, rather than Fodor's (1981) "methodological solipsism" formulation, which says that the taxonomy of an individual's psychological states should not presuppose any essential reference to that individual's environment. I am not concerned here with differences between the two formulations and shall use both in my discussion. Individualism characterized in either way is not a view specifically or only about mental representation or mental content. (Hence there are individualists, such as Stich (1983, Ch. 8), who are sceptical about mental representation.)

[2] Those who have seen computationalism as offering support for forms of individualism in psychology include Devitt (1990, 1991), Egan (1992), Fodor (1981, 1987), and Segal (1989, 1991). The papers by Egan and Segal offer the most explicit arguments from computationalism to some form of individualism.

Mind, Vol. 103 . 411 . July 1994

(1) Cognitive psychology taxonomically individuates mental states and processes only *qua* computational states and processes.

(2) The computational states and processes that an individual instantiates supervene on the intrinsic, physical states of that individual.

Therefore,

(3) Cognitive psychology individuates only states and processes that supervene on the intrinsic, physical states of the individual who instantiates those states and processes.

While the "only" in (1) gives this argument the necessary strength for its conclusion, this will make (1) seem implausibly strong to many who adopt a more pluralistic view of psychological taxonomy. Although I think that individualism in psychology as it has been articulated and defended by some of its leading proponents is adequately expressed by (3) above, and so *any* argument from computationalism to individualism so conceived requires a strong, exclusionary premise such as (1), I do not wish to defend this claim here. In fact, make (1) (and so (3)) as weak or as qualified as you like. Still, the computational argument for individualism should be rejected because its perhaps innocent-sounding second premise is false, and *it* is false because of the possibility of wide computationalism in psychology.

2. *The possibility of wide computationalism*

Suppose that cognitive processing is computational, at least from the point of view of those seeking systematic, scientific, psychological explanations. The states (and the processes which are the transitions between such states) over which a computational psychology quantifies need not be individualistic because the cognitive system to which they belong could be part of a *wide computational* system. That is, the corresponding computational system could transcend the boundary of the individual and include parts of that individual's environment. If this were so, then the computational states of such a cognitive system would not supervene on the intrinsic, physical states of the individual; likewise, the resulting computational psychology would involve essential reference to the environment beyond the individual. The states and processes of a wide computational system are not taxonomized individualistically.

In this section I will concentrate on explaining the coherence of the idea of wide computationalism, i.e., with defending the *possibility* of wide computationalism. I consolidate this defence by considering two objections to wide computationalism in §4, going on in §5 to identify examples of existing research in computational psychology that can be plausibly understood in terms of wide computationalism.

Wide computational systems are computational systems that are not fully instantiated in any individual. Since they literally extend beyond the boundary of the individual, not all of the states they contain can be taxonomized individualistically. Within a wide computational system much of the processing that takes place may well be instantiated fully within the boundary of the individual, but what makes it a wide system is that not all of the computational processes that make up the system are so instantiated. If there are computational (formal) descriptions of both an organism's environment and its mental states, and causal transitions from the former to the latter that can be thought of as computations, there is a process beginning in the environment and ending in the organism which can be viewed as a computation, a wide computation.

To some, the coherence of wide computationalism, its mere possibility, will seem unproblematic. For example, in responding to Martin Davies' claim that "cognitive psychology treats information processing systems (modules) and whole creatures *qua* embedded in particular larger systems and ultimately particular environments" (Davies 1991, p. 482) Gabriel Segal says

> the supervenience base of a representation's content is some larger system in which the representation is embedded. This could be: the whole creature plus its environment, the whole creature, the largest module in which the representation occurs, a sub-processor of that module, a sub-sub-processor of that module, a sub-sub-sub... . Individualism is the thesis that the representational states of a system are determined by intrinsic properties of that system. It seems likely that whole subjects (or whole brains) make up large, integrated, computational systems. Whole subjects plus embedding environments do not make up integrated, computational systems. That is one reason why individualists draw the line where they do: the whole subject is the largest acceptable candidate for the supervenience base because it is the largest integrated system available. (1991, p. 492 ellipsis in original)[3]

Here Segal seems to be conceding the coherence of wide computationalism, claiming that, as a matter of fact, we don't find computational, cognitive systems larger than the individual. This passage identifies precisely where a proponent of wide computationalism disagrees with the proponent of the computational argument: she rejects the claim that the "whole subject", the individual, is, as a matter of fact, "the largest integrated physical system available" for computational, psychological explanation.

Given the coherence of wide computationalism implicit in this passage, it is not surprising that in Segal's surrounding discussion he notes that the disagreement here is properly resolved by an examination of empirical research in computational psychology. Segal himself thinks that the crucial claim that "whole subjects plus embedding environments do not make up integrated, computational systems, can be defended on a posteriori grounds. One would expect, then, an individualist of Segal's persuasion also to consider (2) in the computational argument to have an a posteriori justification, one which while allowing for the mere

[3] Segal characterizes individualism here as a thesis about mental representation, although individualists have typically defended a view about psychological states in general (see footnote 1). Individualists also adopt a more specific view of the nature of the subvenient base than Segal suggests: it is constituted by the intrinsic, physical properties of *the individual*. As we will see, one cannot simply equate individual and computational system, or assume that the latter will be part of the former.

possibility of wide computational systems shows why *our* computational, cognitive systems are individualistic.

Not all individualists adopt this view of (2). For example, Frances Egan (1992) has argued that computational taxonomies are individualistic of their nature: there is something general about taxonomy in computational psychology, or perhaps about computational theory more generally still, which entails that cognitive, computational states and processes are individualistic. If Egan is right, then wide computationalism is inconsistent with some more general feature of computational psychology or computational theory, and (2) is not something which simply happens to be true of the computational systems that we instantiate; rather, it says something true about computational systems per se, and can be defended on a priori grounds.[4]

To bring out the contrast between these two types of defence of (2), and to see why the more a priori defence is problematic in this context, consider the details of Egan's argument. It begins with the claim that the goal of computational theories of cognition is "to characterize the mechanisms underlying our various cognitive capacities" (pp. 444-445).[5] And such theories "construe cognitive processes as formal operations defined over symbol structures" (p. 446). Now,

> Symbols are just functionally characterized objects whose individuation conditions are specified by a *realization function* f_R which maps equivalence classes of physical features of a system to what we might call "symbolic" features. Formal operations are just those physical operations that are differentially sensitive to the aspects of symbolic expressions that under the realization function f_R are specified as symbolic features. The mapping f_R allows a causal sequence of physical state transitions to be interpreted as a *computation*.
>
> Given this method of individuating computational states, two systems performing the same operations over the same symbol structures are computationally indistinguishable. (p. 446).

From this, claims Egan, it follows that "if two systems are molecular duplicates then they are computational duplicates. Computational descriptions are individualistic: they type-individuate states without reference to the subject's environment or social context" (p. 446).

Egan's final conclusion here does not follow, unless one equates computational systems with *subjects*, i.e., with individuals. Yet doing so would beg the

[4] Two points about calling this position Egan's: (a) as we will see, there are strands to Egan's discussion which suggest that she accepts a view of the relationship between computationalism and individualism closer to Segal's; (b) the position is one which I think many individualist find appealing. My concern here is to identify an individualistic position alternative to Segal's; I return to discuss Segal's view in §5 below.

[5] Egan continues: "this goal is best served by theories which taxonomize states individualistically" (p. 445). This may suggest that she sees computational psychology as individualistic for instrumental or pragmatic reasons rather than because of the nature of computational individuation. But, as I hope will be clear from what follows, her actual argument does not appeal at all to whether individualistic or wide taxonomies "best serve" the goal she has identified; rather, it claims that individualism is strictly implied by the *method* by which computational states are individuated,

question against the wide computationalist, for the wide computationalist endorses precisely the claim that there can be computational systems which extend beyond the boundary of the individual. There is nothing in the method of computational individuation itself to which Egan points which implies that the class of physical features mapped by a realization function cannot include members that are part of the environment of the individual. This being so, Egan has not provided a sound argument for why individualism (about computational psychology) follows from the very nature of computational psychology, and so her view does not point to some internal incoherence in the idea of wide computationalism.

Wide computationalism is analogous to wide *functionalism*, the view that the conceptual role defining mental states extends into the world (Harman 1987, 1988; Kitcher 1991). Yet wide computationalism is both more modest and more radical than wide functionalism, and provides the basis for a stronger case against individualism. It is more modest in that it concedes that individualism is true of at least some mental processes and rejects only its all-encompassing nature; it is more radical because it denies something about the notion of a formal or computational system—that it be instantiated in an individual—which is almost without exception taken for granted by individualists, and so undermines the computational argument for individualism in a fundamental way. And it is a more decisive objection to individualism, supposing the "radical" claim to be established, because it not only removes computationalism as one of the major supports of individualism without rejecting computationalism but also provides the basis for arguing from computationalism to a distinctly non-individualistic view of computational psychology itself.

The challenge to the computational argument is not posed by directly defending the claim that psychological states require a broad construal but, rather, by arguing that *the formal* or *computational systems* in which they are instantiated or of which they are a part extend beyond the individual. The distinction between an individual and a cognitive, computational system is central to an understanding of wide computationalism. Even if one thinks that many computational, cognitive systems are fully instantiated in the individual, wide computationalism is a possibility because the boundaries of the individual and those of the computational, cognitive system need not be identical.

An example of a possible wide computational process is the familiar process of multiplication.[6] Typically, apart from multiplication problems that are included in one's "times table", one multiplies numbers by storing intermediate solutions in some written form, usually on paper, and then solving the next component of the problem, storing the result on paper, and so on. The actual process that one goes through in multiplying numbers together typically involves the storage of symbols *on paper*. The problem solving activity itself need not and does not take place solely in one's head; it involves, rather, the use of symbols (and conventions) which are not stored exclusively in the head. A description of the

[6] Thanks to Sydney Shoemaker here; see also Clark (1989, Chs. 4, 7, and 1993, Ch. 6).

process of multiplication must include a description of mathematical symbols, and for most human beings such a description presupposes a reference to something external to the individual organism. A crucial part of the process of multiplication, namely, the storage of mathematical symbols, extends beyond the boundary of the individual. Considered as multipliers, we are part of wide computational systems.

To show the coherence of wide computationalism, this need only be taken as an account of a possible computational, cognitive process, perhaps not one that *we* instantiate. Yet I have described the example in terms of our cognitive processing because I think that human mathematical problem solving, as well as much problem solving more generally, involves (indeed, *essentially* involves) the exploitation of representations in one's environment. The more complex the computational process we engage in—for example, non-trivial mathematical proofs—the more plausible this stronger claim is. Proofs of complex theorems in quantificational logic are rarely carried out entirely in one's head: at least *some* of the symbols are stored externally. What are stored are *pointers* to the symbols that one uses, and while such pointers may be stored inside the head, the symbols to which they point are not stored internally at all: that is why one *needs* a blackboard, pen and paper, or even a calculator.

Not only can a case be made for conceiving of mathematical and logical processes as wide computational processes: the same is true of perceptual and behavioural processes. Wide computationalism is appropriate in cases in which the interaction between an individual and something external to that individual is a crucial part of the computational process being described as an *explanans* in psychology. In the case of perception, it is an intrinsic part of that process that the system accept input from the environment and process it so that further mental processing can proceed. The perceptual process involves an interaction between an individual and her environment. This is in no way incompatible with providing a *computational* account of perception (see §5 below).

Since perception is a process which begins with environmental inputs, inputs which themselves may have a formal description and so be accessible to a computational, cognitive system, all components of the perceptual process can be described as part of a wide computational system. An individualist may object that this characterization of the process of perception simply begs the question. The relevant objects of perception are not external but internal to an individual; for example they are 2-D retinal images, not some type of environmental input. A wide computational account of perception presupposes a view of perception which an individualist should reject.

This objection in effect concedes a weak or negative point I want to make in this section, namely, that the formal or computational nature of mental processing *itself* doesn't entail individualism: one also needs to make a substantial claim about, for example, the objects of perception in order to derive individualism from computationalism. The same is true for any area of cognition which is claimed to be computational. The formality of cognition itself does not entail individualism.

Insofar as this points to a gap between computationalism and individualism in psychology, it allows for the *possibility* of wide computationalism. But a stronger claim about wide computationalism can be formulated and, I think, ultimately defended. Psychological states are computational only insofar as they are part of an implemented formal *system*. (For those who find this controversial, see my discussion in the next section.) But the formal systems of which at least some psychological states are a part are not fully instantiated in any natural individual, i.e., in an organism. So, qua computational states, such psychological states are not instantiated in any individual. Stated in this way, the argument allows one to draw not only the conclusion that a wide computational psychology is possible, but, assuming the truth of computationalism, the conclusion that, for at least some psychological states, such a psychology is *necessary*.

Thus far I have said little about the central notion of *formality*; to further demystify wide computationalism I turn to discuss this notion more explicitly.

3. The notion of formality

Computationalism is sometimes expressed as the view that, since cognition is *formal*, cognitive psychology should be restricted to positing and quantifying over the formal properties of mental states. This expression of computationalism, what Fodor (1981, pp. 226-228) has called the *formality condition*, may make the argument from computationalism to individualism in psychology appear compelling, for the formal properties that mental states have are often thought of as *intrinsic* properties of mental symbols, such as their *shape* and *size*.[7] This conception of computationalism allows one to think of formal properties as a particular species of mental *causal powers*, properties which supervene on the intrinsic, physical properties of the individual, and makes it tempting to view computationalism as providing a general theoretical framework for further specifying the nature of such powers. The task of a cognitive psychology which presupposes computationalism, on this conception of formality, is to discover the intrinsic properties of tokens in the language of thought. In senses that I shall explain below, such properties are both *non-semantic* and *non-physical*.[8]

It should be emphasized first, however, that the formality condition is an *interpretation* of computationalism, or a claim about what the acceptance of compu-

[7] As Fodor says, "formal operations apply in terms of the, as it were, shapes of the objects in their domains" (1981, p. 227, footnote omitted). It is unclear how literally Fodor intends this suggestion, though he relies on it in many places, including more recent work; for example, see his 1987, p. 18.

[8] I thank Robert Stalnaker for emphasizing this point in discussion. The formality condition focuses on only the first of these contrasts, the contrast between formal and semantic properties.

tationalism entails or involves, not simply a statement of computationalism itself. While the notion of formality is often used in computational theory, talk of formal properties as intrinsic *properties* of the individual components of computational systems is, in certain respects, misleading. The conception of formality used in logic, mathematics and computer science, the disciplines which provide the ultimate foundations for computationalism in psychology, is quite distinct from that expressed by the formality condition. In these disciplines the focus is on the properties and behaviour of formal *systems*. A formal system consists of primitives, formation rules, formulae, axioms, and rules of inference. The foundations of logic is concerned, in part, with the relationship between the notions of a formal system, an effective procedure, an algorithm, a computation, and the set of recursive functions. On this conception of formality, what I shall call the *systemic conception of formality*, a given formal system could be expressed in alternative notations and, in principle, could be realized by a nation of people related to each other as the rules of the system specify (cf. Block 1990, Searle 1981). In this sense, the intrinsic, physical properties of symbols in a formal system are arbitrary.[9]

On the systemic conception of formality, there is little talk of the formal properties that *particular* symbols have. The sorts of "formal properties" which are primarily discussed, properties such as being closed under modus ponens, being transitive, being compact, and being sound and complete, are properties of formal systems or, derivatively, properties of symbols as elements of formal systems. Computational processes, operations and instructions are often thought of as formal, but this is to say only that they can be adequately described as the result of the application of rules or algorithms which constitute the system to which they belong. Insofar as particular symbols in a formal system have formal properties, it is not clear whether such properties are intrinsic or extrinsic properties of the symbols themselves. For example, an instance of the symbol "A" will lead to an instance of the symbol "B" and do so in virtue of its "shape" in a formal system containing only the rule "A $\Rightarrow$ B". But since this formal property, having that particular shape, has that effect only in a formal system with a rule of that type, one should be wary of identifying such formal properties with intrinsic causal powers that symbols possess. In any case, what is clear is that such properties have the causal significance that they do only insofar as the symbols to which they are attributed are part of a formal system.

The systemic conception of formality, which I shall rely on in the remainder of the paper, makes it natural to express computationalism as the view that cognitive psychology ought to be pitched at a computational *level of description*. I said above that the formal properties of mental states are supposed to be both non-semantic and non-physical, and I want to explain what these two contrasts imply

[9] Interestingly, Rollins (1989) and Devitt (1990, 1991) both draw something like the distinction I am drawing here between the formality condition and the systemic conception of formality, although neither seems to think the formality condition misleading in the way in which I am claiming it is.

about the (narrow) computationalist's conception of cognition by looking at the two different conditions that a computational level of description of psychological states and processes must satisfy.

In contrast to the *physical* level of description, the computational level is distinct from and irreducible to the levels of description which characterize the physical realizations of a particular formal system. The same computational program, the same formal system, can be instantiated in many physically distinct ways. Given computationalism, this is the sense in which psychology is autonomous of the physical sciences. It is this autonomy, and so the contrast between the formal and the physical, which, I think, underlies the first premise of the computational argument.

In contrast to the *semantic* or representational level of description, the computational level specifies the properties of mental symbols and the rules constituting the formal system of which those mental tokens are a part without reference to what, if anything, those symbols represent. The proponent of the computational argument for individualism claims that, perhaps unlike the semantic level of description, the computational level specifies properties which are determined by the intrinsic, physical states of the organism in which they are instantiated.[10] This feature of the properties specified at the formal level of description makes the second premise of the computational argument intuitively plausible.

4. Two objections to wide computationalism

One prima facie strength of wide computationalism is that it is fairly non-committal regarding the precise computational character of cognition. For example, it would seem to be compatible with both "classical" and connectionist conceptions of computationalism in psychology. Yet this potential strength of wide computationalism may be seen as its Achilles' heel by someone pressing the issue of the degree to which wide computationalism is a *realist* view of computational psychology: to what extent does the plausibility or even the possibility of wide computationalism turn on a view of computationalism that is committed to little more than the utility of the computational *metaphor* in psychology? To put it slightly differently: does wide computationalism presuppose that computational explanations in psychology only *model* the phenomena they purport to explain, in the same way that there are computational models of other phenomena, such as the motions of planetary systems? If so, then wide computationalism will be a view of little significance for computational *psychology*. Central to the computa-

[10] This contrast between computational and intentional taxonomies is clear in Egan (1992), where a defence of individualism about computational psychology is combined with a defence of wide intentional content. One could, I suppose, hold just the reverse of this: that computational taxonomies may be wide but intentional taxonomies must be narrow. Exploring this option should be of interest to those who accept wide computationalism but wish to defend an individualistic view of intentionality.

tional paradigm in psychology is the idea that an individual's mind is not simply described or modelled by a computer program; cognition is *rule-guided*, not simply *regular* (Bennett 1989). Wide computationalism is possible only if one relies on a weak reading of the computational metaphor, a reading which does not do justice to computationalist commitments in contemporary cognitive psychology.

To understand how a wide computational system could produce rule-guided behaviour, consider how a *narrow* computational system could do so. Since standard personal computers are paradigm cases of narrow computational systems, we can make our discussion more concrete by asking how *they* produce behaviour by actually following rules. Computers follow rules by instantiating or implementing programs constituted by such rules. So what is it to implement a program? For a physical device to be *capable* of implementing a given program is for it to have its physical states configured in such a way that transitions between those states are isomorphic to transitions between states that the program specifies, i.e., there is a mapping from equivalence classes of physical states to the symbolic states that constitute the program. Since implementational power is characterized in terms of the mathematical notion of isomorphism, there is a large number of actual programs and an infinite number of possible programs which any given physical device can implement. One closes the gap between the power to implement and actual implementation by identifying the appropriate causal interaction between the physical storehouse for the program (e.g., a physical disk) and the computer itself. So, in response to the grand epistemological, scepticism-mongering question, "Of the infinite number of programs that a computer could be implementing, how do you know that it is implementing *this* program?", we say: "It implements this one because it is this one that is encoded on the disk we inserted." (And since a physical disk is simply one type of storehouse for a program, we could replace reference to a physical disk here by reference to anything else a program is stored on.)

This view of implementation may make it sound as though the program is *epiphenomenal* to the physical operation of the computer, raising doubts about it as an account of rule-guided behaviour: in what sense is the behaviour that the computer generates anything more than *regular* behaviour, behaviour that appears to be rule-governed but, in fact, is not? I should make it clear that I think that the program *does* play a causal role in the behaviour of the physical device, and that the behaviour it produces is thus rule-governed and not merely regular. But while we may wish to say that the machine behaves in the way it does *because* of how it was programmed (i.e., because of the program it instantiates), we should be sure to distinguish this sort of "downward" causation from that which exists between the physical states themselves. Unless there is massive causal pre-emption "from above", symbolic states can't be viewed as the direct causal antecedents (the efficient causes) of later physical states. In general, to understand the causal role that higher-level states play in the production of behaviour, we need a broader conception of the notion of a causal role than is typically assumed (see Wilson 1993, 1994, in press).

Although I have stated this view of how computers produce rule-guided behaviour in terms of familiar narrow computational systems, the narrowness of the system plays no significant role in the view; much the same story can be told of a *wide* computational system. The account of implementational power is precisely the same: the wide computational system has the power to implement just those programs for which there is an isomorphism between the system's physical states and the symbolic states the program specifies. The account of actual implementation is a generalization of that in the case of narrow computational systems: a wide computational system implements the "program" physically stored in the environment with which it causally interacts. Determining the proper symbolic description of aspects of an organism's environment is an a posteriori matter, much as doing so with respect to an organism's *internal* structure is.

"Program" occurs in scare-quotes here because of two important differences between the programs that run on standard computers and those that (narrow or wide) computationalists claim run on us: (i) unlike the programs that *we* encode on physical disks, precisely what symbolic interpretations can be given either to aspects of an organism's environment or to its internal structure (or both) are things that must be *discovered*; (ii) these interpretations may not turn out to be elaborate enough themselves to warrant the label "program". Significantly, (i) and (ii) distinguish what we know (and love?) as actual computers from organisms. We simply are not in the appropriate epistemological position to claim either that our brains or our brains plus our environments instantiate programs in precisely the sense that computers do. And in light of the similarities and differences between us and computers that emerge from empirical research, we will be able to decide whether "programs" or "internal languages" are appropriate categories with which to develop psychological explanations. None of this involves adopting a weak understanding of the computational metaphor in psychology, only some epistemic caution that should be adopted whether one defends narrow or wide computationalism.

The idea that by going wide one gives up on something crucial about computationalism reflects a deeply Cartesian view of the mind, a vestige of thinking of the mind and body as distinct substances, which survives within contemporary materialist and naturalistic views of the mind. This vestige is the idea that there is something special about the mind, about what is "in the head", that justifies the ascription of computational states to it, which is not shared with extra-cranial reality; there is a bifurcation between mind and mere matter which makes *only* narrow computationalism a serious option within psychology (cf. Segal 1991, p. 492, quoted above). I shall refer to this idea as *Cartesian computationalism* and will say more about it in my conclusion.

Let me turn to a second objection to wide computationalism, one which introduces broadly empirical grounds for doubting that we are wide computational systems. As Egan (1992, pp. 446, 457) notes, citing examples of research in early vision and in syntactic and morphological analysis in linguistics, the psychological processes for which there are the most satisfying computational accounts are *modular*: they are domain-specific and informationally encapsulated.[11] That is, we have had our greatest empirical successes in computational psychology in explaining the character of psychological processes which function with relative independence from even much of the *internal* workings of the cognitive system, let alone the external environment of the individual. If empirical success has come within computational psychology only or even predominantly with the correctness of the presumption of modularity, then that should cast doubt on the idea of developing an empirically adequate *wide* computational psychology.

Fodor (1983) makes this point about the relationship between modularity and computational psychology more poignantly by arguing that "global systems are per se bad domains for computational models" (p. 128). Specifically, what he calls *central* processes, such as problem solving and belief-fixation, are unlikely to have computational models precisely because they are, in his view, non-modular. The non-modularity of central processes gives one reason to be sceptical about the real (versus mere) possibility of an adequate computational psychology explaining them. And what is true of central processes, processes which have access to a variety of representational inputs, is also true of wide computational processes, processes which access representations that are *outside* the individual.

Suppose we agree that, by and large, the empirical successes that cognitive science has had thus far have involved highly modular systems, such as those employed in visual perception and phoneme recognition. Perhaps this is for a deep reason, such as its only being highly modular systems which *are* computational; alternatively, it could be due to a relatively shallow reason, such as its being only highly modular computational processes that theorists can readily understand as computational. In either case, there is nothing here that allows for the application of a point about central processes to wide computational processes *since the latter can also be modular.* As I hope the discussion in the next section indicates, contemporary research in cognitive psychology that is properly considered as positing wide computational systems involves highly modular systems.[12]

The implicit premise in the argument from modularity to individualism sketched above—that modular systems are taxonomized individualistically—is false because modular systems may well encapsulate information that is in the

[11] The notion of modularity at work here is that articulated by Fodor (1983). While Fodor lists other features of modular systems (e.g., they are fast and mandatory), domain-specificity and informational encapsulation are the two most central. Although these notions themselves warrant some conceptual elaboration (see Wilson, in press), let me here simply provide an intuitive gloss on each. A *domain-specific* system is one which operates on some particular type of information (a domain). An *informationally encapsulated* cognitive system acts as an input-output function on a specifiable and specific set of informational inputs and outputs: it encapsulates some kind or kinds of information, and is insensitive to other information.

[12] In fact, I think that there is no conceptual problem in even central processes, such as inference, being modular: even *they* may be domain-specific and informationally encapsulated. Dan Sperber (1994) has argued for this view, and I take the work of Cosmides and Tooby in evolutionary psychology to provide empirical support for it; see, for example, Cosmides (1989) and Cosmides and Tooby (1987).

individual's *environment*, not elsewhere in the individual. Thus, the module may be a part of a computational system all right, but a *wide* computational system. Neither of the chief two features of modular cognitive systems, their domain-specificity and informational encapsulation, implies that such systems cannot be properly viewed as parts of wide computational systems. Being wide and being modular are compatible properties for a cognitive system to possess—and so modularity does not entail individualism—because the location of the information with respect to which that system is encapsulated does not affect that system's domain-specificity. The real question to be answered is this: are narrow computational accounts of given modular systems always *explanatorily richer* than their wide rivals? It is only if the answer to this question is "Yes" that wide computationalism can be rejected as less plausible in general than narrow computationalism.

5. Wide computationalism in cognitive psychology

Although the possibility of wide computationalism suffices to show that the second premise of the computational argument is false, for those antecedently disposed to think that wide computationalism is coherent the real interest in the computational argument lies in the claim that *we* are plausibly seen as wide computational systems. I think that wide computationalism *is* made plausible by recent computational research in both human and animal cognition. Showing wide computationalism to be not only a coherent but a plausible view of our cognitive processing would both consolidate and broaden my objection to the computational argument. I shall discuss two examples of research in cognitive psychology which show wide computationalism in action.[13]

Sekuler and Blake (1990) devote a significant section of their chapter on spatial vision and form perception to a discussion of an approach to form perception pioneered in the work of Campbell and Robson (1968), an approach known as *multiple spatial channels theory.* The basic idea of the approach is that there are specific stimuli which individual sets of neurons are sensitive to, these stimuli being decomposable into sinusoidal *gratings*. These gratings are relatively simple, having only four relevant parameters: spatial frequency, contrast, orientation, and spatial phase. Any figure composed of these gratings is definable formally in terms of these four parameters. The bold and controversial claim of this research program is that *any* natural scene in an organism's environment can be decomposed into its gratings, and this fact explains a great deal of human form perception, including its limitations.

On this conception of form perception, part of the task of the perceptual psychologist is to identify formal primitives that adequately describe the visual envi-

[13] Thanks to both David Field and Frank Keil for useful discussion of the material in this section; they should not, however, be saddled with the conclusions I draw here.

ronment, and to specify algorithms which apply to these primitives to determine complete visual scenes. To see what this means, take a case simpler than human vision, that of a lens projecting an image of an object onto a piece of white paper. Figure 1 shows the *transfer function* for two lenses, which plots how contrast is transferred through the lens from object to image, and is defined over a range of spatial frequencies. As input, it takes contrast in an object, producing as output contrast in the image. We can likewise define a *contrast sensitivity function* for the human visual system, which takes the same inputs from the world to produce a visual output (see Figure 2). The formal system that perceptual psychologists working within this paradigm study is not instantiated in any individual: it includes but is not restricted to the intrinsic properties of an individual. This is reflected in the actual methodology employed by such psychologists, which involves the extensive and complex mathematical analysis of natural scenes into their computational primitives. Such analysis appears to be an intrinsic part of the multiple spatial channels paradigm, not simply something preliminary to real perceptual psychology.

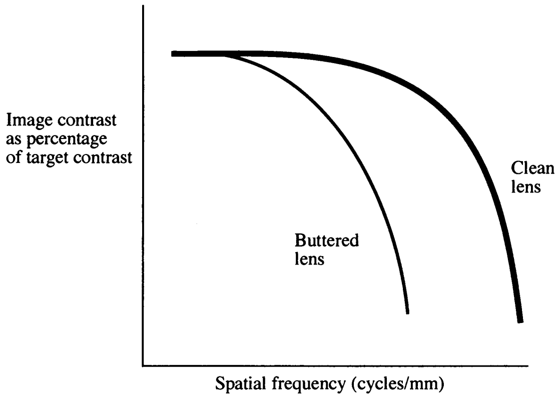

Figure 1 Two transfer functions for a lens. The curves specify how contrast in the image formed by the lens is related to contrast in the object.
(Both figures reproduced from Sekuler, R. and Blake, R. 1990: *Perception*, with permission of McGraw Hill.)

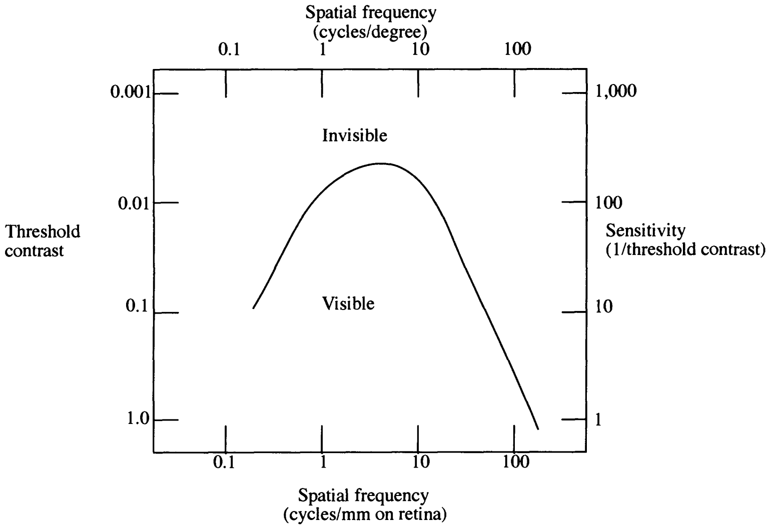

Figure 2 A contrast sensitivity function for an adult human. The upper horizontal axis is scaled in units specifying the number of paris of light and dark bars of the grating falling within one degree of visual angles on the retina.

Gallistel (1989a) reports research on the conceptions of space, time and number that a variety of animals have, including bees, rats, and ants. One of Gallistel's primary conclusions is that purely sensory-based models of a range of animal behaviour are inadequate. Rather, the evidence overwhelmingly suggests that these animals construct quite complex representations of their environments and use these to guide their behaviour. Gallistel argues that such representations are computational, that there is strong evidence that these animals instantiate modules which are sensitive to the formal (e.g., the geometric) structure of their environments, and that this sensitivity is responsible for their navigation through their physical environments. For example, in ants and bees the computational process of *dead reckoning* (which integrates velocity with respect to time) takes as inputs the animal's solar heading, forward speed, and a representation of the solar azi muth, producing as output a representation of the creature's position relative to some landmark, such as a nest. The *ephemeris function*, which produces the third of these inputs, takes as its inputs a sighting of the sun and time on some endogenous clock (see Gallistel 1989b, pp. 70-76). In both of these cases, the computational process extends beyond the boundary of the individual.

Gallistel calls his view a *computational representational* perspective on animal cognition. Of animal navigation, Gallistel says:

> routine animal movements are governed by a navigational process closely analogous to everyday marine practice. This practice rests on an extensive isomorphism between the geometry of motion and position and the computational processes that underlie navigation. At the neurophysiological level of analysis, the hypothesis implies that the mathematical description of the processes in the animal brain that function during animal navigation parallels the mathematical description of the computations a human or computerized navigation system makes. (1989a, pp. 176-177)

This quotation suggests that Gallistel, like those working in the multi-channels paradigm in form perception, does not see anything mysterious in positing an extensive isomorphism between the formally described properties of an environment and those of mental processes.[14]

Two central postulates of these otherwise diverse research programmes is that the environment of the organism has a certain formal structure to it, and that the organism's sensitivity to this structure explains core parts of its cognitive performance. Characterizing the specific nature of the environment in computational terms appears to be a central part of the implicit conception of cognitive psychology in these research programmes. Despite some of Gallistel's own claims about the view he advocates (see below), I see no way in which this is merely additional or peripheral to these research programmes, and in the remainder of this section I shall defend the view that both research programmes support the view that we and our biological kin are parts of wide computational systems.

Recall Segal's view of the relationship between computationalism and individualism, discussed in §2: that, as it turns out, our cognitive systems are narrow, not wide, computational systems. Thus, wide computationalism should be rejected because although it is a coherent view there is *no* research which, in Segal's words, treats "whole subjects plus embedding environments" as "integrated,

[14] Cf. Gallistel's claim that the isomorphism between computational processes instantiated in the head and certain "formal properties" in the environment is responsible for the successful navigation behaviour that many animals exhibit: "there is a rich formal correspondence between processes and relations in the environment and the operations the brain performs. Brain processes and relations recapitulate world processes and relations. The recapitulation is not fortuitous. To fit behaviour to the environment, the brain creates models of the behaviour-relevant aspects of the environment. the formal properties of the processes that realize these models reflect the formal properties of the corresponding external reality because these processes have been subject to evolution by natural selection. Selection has been based on the fidelity of these processes to the external reality. Evolution by natural selection creates rich functioning isomorphisms between brain processes and the environment, and learning is to be understood in terms of these isomorphisms (1989b, p. 27)."

computational systems". This is a strong empirical claim, one to which in this section I have provided two prima facie counter-examples. The individualist who wishes to defend Segal's claim needs to explain away the appearances. Once the coherence of wide computationalism is conceded—its mere possibility—there are likely to be many such appearances to explain away, and so such an individualist faces a prima facie difficult task.

I think that the most promising way to defend the computational argument from the line of objection I have developed is to concede that even if one *can* view the individual as embedded within a wide computational system, there is no *explanatory* motivation for doing so. For we can also view the individual itself (or, more accurately, a *part of* that individual) as a computational system, a narrow computational system, and *doing so is always adequate for computational, psychological explanation.* Consider the two examples that I have given of purportedly wide computational research. Even if the multi-channels paradigm does seem to posit a formal structure to the environment, it could also be viewed as claiming that the retinal image has such a structure. If this view of the paradigm is correct, then while one *could* view the paradigm as being computationally wide, there is no need to adopt this view of it. Likewise, while the computational representational view that Gallistel defends might seem to view an individual as part of a wide computational system, a system which includes features of that individual's environment, one could also see his view as positing an interesting isomorphism between *two* formal systems, one of which is fully instantiated in the individual. Crucially, cognitive psychology is the study of this *narrow* computational system. Since the form of this defence of the computational argument is the same in each of these two cases, I shall develop it and respond to it by focusing only on Gallistel's view.

A general feature of Gallistel's view of animal cognition that might be thought difficult to reconcile with my interpretation of it is that much animal behaviour is governed by internal maps and mathematical representations of the environment rather than direct sensory input. This view ascribes to an animal a high degree of autonomy from its environment, and this aspect to Gallistel's view at least *sounds* individualistic: animals navigate, for example, by *internal maps*, not by sensory tracing, homing, or other environmentally interactive methods. In characterizing how an animal navigates, we abstract away from its actual environment and concentrate on the intrinsic features of its map or model of that environment.

Several of Gallistel's comments about his own project offer support for this type of individualistic interpretation. Gallistel says that his "agenda is a reductionist one: to understand learning as a neuronal phenomenon" (1989b, p. 24), going on to say that studying the total [wide] computational system is simply a "necessary prelude to understanding what the system does in terms of what its elements do" (1989b, p. 24). Figuring out the computational structure to an organism's environment, while *methodologically* necessary, is peripheral to an understanding of the nature of learning itself, and suggests that there is no deep sense in which Gallistel advocates a view of learning as a wide computational process. In addition, in the conclusion to his book Gallistel says that "[t]he structure of the computational mechanisms is dictated by the formal structure of the representations to be computed and by the sensory or mnemonic data from which they are computed" (1989b, p. 581), suggesting that he sees the *computational* system of interest to the psychologist as fully instantiated in the individual organism.

One response to this interpretation corresponds to and reinforces the weak or negative claim that I made in §2. Even if one *can* see Gallistel's view as a narrow computational approach, the fact that it also can be given a wide computational interpretation shows that computationalism itself does not entail individualism: computational systems need not be individualistic. Yet this response does not address the issue of which of these interpretations of Gallistel's view has greater explanatory adequacy. A second response, corresponding to one of the stronger claims made in §2—about the plausibility or even necessity of wide computationalism in psychology—addresses this issue.

There are two aspects to this second response, one elaborating on the concepts of explanation and explanatory adequacy, the other concentrating on the application of these concepts so understood to particular examples within computational, cognitive psychology. Let me say something brief about the first of these aspects. Were the individualist to claim, in effect, that there is no explanatory need to posit wide computational systems because whenever we can do so we can also identify a corresponding, narrow computational system,[15] she would be presupposing *some* conception of explanatory adequacy, although one yet to be explicitly articulated. Elsewhere (Wilson 1994) I have developed a concept of explanatory adequacy in terms of the notions of *causal depth* and *theoretical appropriateness* and argued that wide psychological explanations are sometimes causally deeper and more theoretically appropriate than individualistic psychological explanations. I think that this argument can be adapted to apply specifically to show that wide *computational* explanations are sometimes more explanatorily adequate than narrow computational explanations in psychology, though I shall not argue this point here. What I want to do instead is make several points pertinent to the second aspect of the issue of explanatory adequacy, points about the research reported by Gallistel discussed above.

The sorts of behaviour that Gallistel is trying to explain involve a flow of information from the environment to the organism. One of Gallistel's main points is that this is not a constant flow of information, as suggested by simple sensory

[15] This claim is very similar to one that proponents of the narrow content program in psychology have sometimes made: that whatever explanatory work a wide taxonomy of psychological states (e.g., that of folk psychology) does can also be done by a corresponding narrow taxonomy of those states. This claim has sometimes been expressed as the view that wide taxonomies, insofar as they are explanatory, can be factored into narrow taxonomies, plus some external remainder, in much the way that the concept of weight can be factored into the concept of mass, plus that of gravitational force (see Field 1981; Fodor 1987, Ch. 2). I have argued against such views (1992, 1993, in press, forthcoming).

accounts. Yet it would be a mistake to think of his view as denying that (formal) properties of the environment play any significant role in a complete, cognitive explanation of animal navigation.

First, even if an animal's behaviour in navigation is primarily governed by internal maps, these maps are updated by periodically acquired information about the organism's movement and the relative position of objects in the environment. An account of these updates ("fix-taking", in Gallistel's terms) is a necessary part of a complete psychological explanation of the behaviour. If one takes the computational representational view to provide a comprehensive paradigm for the investigation of animal learning and cognition, then fix-taking must be accommodated within that paradigm. The narrow computationalist view allows that relations between psychological states themselves can be computational. The wide computationalist proposes a natural extension of this view to allow organism-environment interactions, such as fix-taking, to be subject to a (wide) computational approach. The individualist must explain processes such as fix-taking in some other way, since her claim is that computational systems are fully instantiated in individuals. The wide computationalist is able to offer a view of fix-taking that has greater explanatory unity than that available to the individualist.

Second, as shown by the above examples of environmental inputs in Gallistel's account, characterizing cognitive, computational states *as representations* sometimes requires non-individualistic descriptions. If a representational psychology may violate individualism in the descriptions it offers of psychological states, then why must a computational psychology be individualistic? This question is not rhetorical, for (i) whether representational psychology violates individualism in particular cases is a substantial question, and (ii) the width of representational content might be thought compatible with the narrowness of computational psychology because of the different functions that intentional and computational ascriptions serve. Egan (1992) has argued that the role of semantics in computational psychology is to provide explanatory *models*, models which may be either narrow or wide, for individualistic computations. If Egan is right about the different explanatory roles that content and computational ascriptions play, then this provides a principled reason for distinguishing between computational and representational aspects to psychology with respect to individualism. To discuss this issue further one would need to re-examine the often-invoked trichotomy between the physical, the syntactic (formal, computational), and the semantic (representational). My own view is that it is most plausible to see the computational and representational levels of description as playing very similar explanatory roles in psychology. Again, however, the issue here cannot be suitably resolved without some discussion of the notion of explanatory adequacy and its application to computational psychology.

6. Conclusion

The computational argument for individualism should be rejected because its second premise, the assumption that computational processes in general are individualistic, is false in light of the possibility and plausibility of wide computationalism in cognitive psychology. Much of my discussion of the second premise of the computational argument aims to explain the coherence of wide computationalism, what I have been calling its mere possibility. I think, however, that the most interesting issue concerns not the coherence of wide computationalism but the extent to which a wide computational research strategy is and could be employed within cognitive psychology. Given that, why have I concentrated on the mere possibility of wide computationalism, rather than its plausibility in cognitive psychology?

The central idea behind wide computationalism is extremely simple. However, fleshing out the idea and being explicit about the implications it has for issues in philosophical psychology allow one to see the respects in which it represents a radical departure from the conception of the mind underlying much contemporary research in computational psychology, what I have called Cartesian computationalism. Precisely because Cartesian computationalism is typically unidentified and unexamined, even the basic idea behind wide computationalism is likely to produce knee-jerk puzzlement. It is for this reason that I have spent so much time in §§2-4 demystifying wide computationalism by articulating what the mere possibility of wide computationalism amounts to.

The more interesting issue of the plausibility of wide computationalism as a perspective on research in cognitive psychology is one whose resolution, as we have seen, turns partially on further discussion of general notions, such as that of explanatory adequacy and causal role, as well as an analysis of notions more specific to computational psychology, such as modularity and formality. Of particular importance is further discussion of the relationship between the notion of formality and *representationalism* in psychology. Such discussion and analysis ought to shed further light on the purported examples of wide computational research that I provided in §5, though I am far from suggesting that the illumination will be unidirectional. The examination of actual research in computational psychology on which the issue of the plausibility of wide computationalism in psychology turns should also facilitate the more conceptual work that remains to be done.

While I think that much contemporary research in computational psychology and (especially) artificial intelligence operates within a Cartesian computational framework, I am not offering an external, a priori critique of such research. Rather, accepting the assumption that psychological states are computational, I am questioning another assumption—that computational, cognitive systems are always completely instantiated in individuals—that usually goes along with it. If I am correct in thinking both that this assumption is typically unquestioned and that it is, in some cases, false, then wide computational psychology would seem worth further investigation. If wide computationalism is not only coherent but a plausible view of at least some existing research within the computational paradigm, then the computational argument can be turned on its head: individualism does not impose a constraint on the individuation of mental states precisely because, in at least some cases, psychological states *are* considered as (wide) computational states for the purposes of psychological explanation.[16]

Department of Philosophy
Queen's University
Kingston
Ontario
K7L 3N6
Canada

ROBERT A. WILSON

REFERENCES

Bennett, Jonathan 1989: *Rationality*. Indianapolis, Indiana: Hackett Publishing.

Block, Ned 1990: "Troubles with Functionalism", excerpted in Lycan 1990, pp. 444-468.

Campbell, F.W. and Robson, J.G. 1968: "Application of Fourier Analysis to the Visibility of Gratings". *Journal of Physiology* 197, pp. 151-166.

Clark, Andy 1989: *Microcognition*. Cambridge, MA: MIT Press.

——1993: *Associative Engines*. Cambridge, MA: MIT Press.

Cosmides, Leda 1989: "The Logic of Social Exchange: Has Natural Selection Shaped How Humans Reason? Studies with the Wason Selection Task". *Cognition* 31, pp. 187-276.

Cosmides, Leda and Tooby, John 1987: "From Evolution to Behaviour: Evolutionary Psychology as the Missing Link", in John Dupré, ed., *The Latest on the Best*. Cambridge, MA: MIT Press, pp. 277-306.

Davies, Martin 1991: "Individualism and Perceptual Content". *Mind* 100, pp. 461-484.

Devitt, Michael 1990: "A Narrow Representational Theory of the Mind", in Lycan 1990, pp. 371-398.

——1991: "Why Fodor Can't Have it Both Ways", in Barry Loewer and George Rey, eds., *Meaning in Mind*. Cambridge MA: Basil Blackwell, pp. 95-118.

Egan, Frances 1992: "Individualism, Computation, and Perceptual Content". *Mind* 101, pp. 443-459.

Field, Hartry 1981: "Mental Representation", in Ned Block, ed., *Readings in Philosophy of Psychology*, vol. 2. Cambridge, MA: Harvard University Press, pp. 78-114.

[16] An early version of this paper was presented at the 1992 meeting of the Canadian Philosophical Association and the Society for Philosophy and Psychology; I would like to thank my commentators on those occasions, Wayne Henry and Brian Smith, respectively, as well as the audiences there for useful feedback. Thanks also to Frances Egan, Sydney Shoemaker, Robert Stalnaker, and J.D. Trout for written comments on earlier drafts as well as helpful discussion; and to an anonymous referee for *Mind*.

Fodor, Jerry 1981: "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology", in his *Representations*. Brighton, Sussex: Harvester Press, pp. 225-253.

——1983: *The Modularity of Mind*. Cambridge, MA: MIT Press.

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

Gallistel, C.R. 1989a: "Animal Cognition: The Representation of Space, Time and Number". *Psychology Annual Reviews* 40, pp. 155-189.

——1989b: *The Organization of Learning*. Cambridge, MA: MIT Press.

Harman, Gilbert 1987: "(Nonsolipsistic) Conceptual Role Semantics", in Ernie LePore, ed., *New Directions in Semantics*. London: Academic Press, pp. 55-81.

——1988: "Wide Functionalism", in Stephen Schiffer and Diane Steele, eds., *Cognition and Representation*. Boulder, CO: Westview Press, pp. 1-12.

Kitcher, Patricia 1991: "Narrow Psychology and Wide Functionalism", in Richard Boyd, Phil Gasper, and J.D. Trout, eds., *The Philosophy of Science*. Cambridge, MA: MIT Press, pp. 671-685.

Lycan, William, ed., 1990: *Mind and Cognition*. Cambridge, MA: Basil Blackwell.

Rollins, Mark 1989: *Mental Imagery*. New Haven, CT: Yale University Press.

Searle, John 1981: "Minds, Brains, and Programs", in John Haugeland, ed., *Mind Design*. Cambridge, MA: MIT Press, pp. 282-306.

Segal, Gabriel 1989: "Seeing What is Not There". *The Philosophical Review* 98, pp. 189-214.

——1991: "Defence of a Reasonable Individualism". *Mind* 100, pp. 485-494.

Sekuler, R. and Blake, R. 1990: *Perception*, 2nd edition. New York: McGraw-Hill.

Sperber, Dan, 1994: "The Modularity of Thought and the Epidemiology of Representations", in Lawrence Hirshfeld and Susan Gelman, eds., *Domain Specificity in Cognition and Culture*. New York: Cambridge University Press.

Stich, Stephen 1983: *From Folk Psychology to Cognitive Science*. Cambridge, MA: MIT Press.

Wilson, R.A. 1992: "Individualism, Causal Powers, and Explanation". *Philosophical Studies* 68, pp. 103-139.

——1993: "Against A Priori Arguments for Individualism. *Pacific Philosophical Quarterly* 74, pp. 60-79.

——1994: "Causal Depth, Theoretical Appropriateness, and Individualism in Psychology". *Philosophy of Science*, 61, pp. 55-75.

——in press: *Cartesian Psychology and Physical Minds*. New York: Cambridge University Press.

——forthcoming: "The Place of Folk Psychology in Explaining Human Behaviour: Computationalism, Individualism, and Narrow Content". *Synthese*.