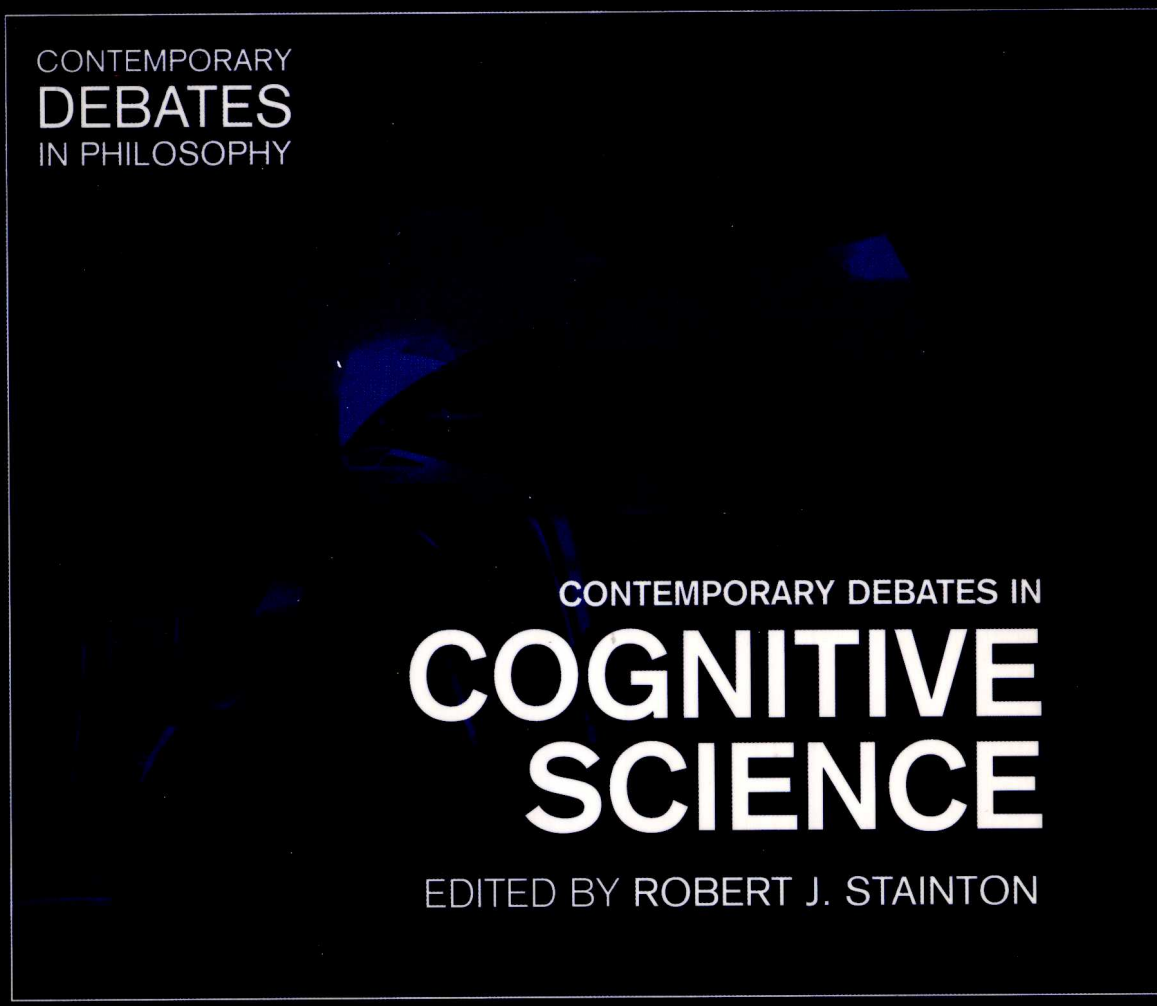
CONTEMPORARY
DEBATES
IN PHILOSOPHY



CONTEMPORARY DEBATES IN

COGNITIVE SCIENCE

EDITED BY ROBERT J. STAINTON

Blackwell
Publishing

Contemporary Debates in Cognitive Science

Edited by

Robert J. Stainton

Blackwell
Publishing

Acknowledgments

A great number of people – too many to mention here – made suggestions of topics, authors, etc. as the volume took shape. I am grateful for their (sometimes anonymous) help. I am also grateful, of course, to the authors for their excellent contributions . . . and for their patience while the book was completed. I would also like to thank the extremely helpful crew at Blackwell. Finally, I must acknowledge the Carleton University Doctoral Program in Cognitive Science. My sense of what cognitive science is, and my understanding of all the issues discussed in this volume, derive largely from a decade or so spent working with colleagues and students in that wonderful unit. Special thanks to Andy Brook, its founder. This book is dedicated to him.

Brie Gertler is Associate Professor of Philosophy at the University of Virginia. She works on issues in the philosophy of mind, including the problem of self-knowledge, the status of physicalism, and the scope of mental content. Her papers have appeared in *Analysis*, *Noûs*, *Philosophical Studies*, *Philosophy and Phenomenological Research*, and other journals. She is editor of *Privileged Access: Philosophical Accounts of Self-Knowledge* (2003).

Gerd Gigerenzer is Director at the Max Planck Institute for Human Development, Berlin; former Professor of Psychology at the University of Chicago; and John M. Olin Distinguished Visiting Professor, School of Law, University of Virginia. His books include *Simple Heuristics that Make Us Smart* (with Peter Todd and the ABC Research Group, 1999), *Adaptive Thinking: Rationality in the Real World* (2000), *Bounded Rationality: The Adaptive Toolbox* (with Reinhard Selten, 2001), and *Reckoning with Risk* (2002, published in the US as *Calculated Risks*, 2002). He is the winner of the Association of American Publishers Prize for the best book in the Social and Behavioral Sciences, the American Association for the Advancement of Science (AAAS) Prize for Behavioral Science Research, and the 2002 German Science Book of the Year Prize.

Terence Horgan is Professor of Philosophy at the University of Arizona. He works, often collaboratively, in philosophy of mind, philosophy of cognitive science, epistemology, decision theory, metaphysics, metaethics, and philosophy of language. He and John Tienson coauthored *Connectionism and the Philosophy of Psychology* (1996). His recent articles include "Phenomenal intentionality and the brain in a vat" (with J. Tienson and G. Graham), in R. Schantz (ed.) *The Externalist Challenge* (2004), "Internal-world skepticism and the self-presentational character of phenomenal consciousness," in U. Kriegel and K. Williford (eds.) *Consciousness and Self-Reference* (forthcoming), and "Some ins and outs of transglobal reliabilism (with D. Henderson), in S. Goldberg (ed.) *The Externalism/Internalism Debate in Semantics and Epistemology* (forthcoming).

Ray Jackendoff is Professor of Philosophy and codirector of the Center for Cognitive Studies at Tufts University. He previously taught at Brandeis University. His research is centered on natural language semantics and its relation to human conceptualization, with major side interests in syntax, social cognition, consciousness, and musical cognition. He is a Fellow of the American Academy of Arts and Sciences and of the American Association for the Advancement of Science. He was President of the Linguistic Society of America in 2003 and of the Society for Philosophy and Psychology in 1991. His most recent books are *Foundations of Language* (2002) and, coauthored with Peter Culicover, *Simpler Syntax* (2005).

Kirk Ludwig is Professor of Philosophy at the University of Florida. He has published articles in the philosophy of language, mind, action, and epistemology. He is editor of *Donald Davidson* (2003) and coauthor with Ernie Lepore of *Donald Davidson: Meaning, Truth, Language and Reality* (2005), and the forthcoming *Donald Davidson: Truth-theoretic Semantics*.

William G. Lycan is William Rand Kenan, Jr. Professor of Philosophy at the University of North Carolina. He is author of *Logical Form in Natural Language* (1984), *Knowing Who* (with Steven Boër, 1985), *Consciousness* (1987), *Judgement and Justification* (1988), *Modality and Meaning* (1994), *Consciousness and Experience* (1996), *Philosophy of Language: A Contemporary Introduction* (2000), and *Real Conditionals* (2001). He has also edited an anthology, *Mind and Cognition* (Blackwell, 1990, 1999).

David Matheson completed his PhD in the Department of Philosophy at Brown University in 2003. His doctoral work, supervised by the prominent epistemologist Ernest Sosa, explored the relation between everyday knowledge commitments and epistemological theory. He is currently teaching in the Department of Philosophy at Carleton University (Ottawa, Canada), and is an executive committee member of the Canadian Society for Epistemology.

Robert J. Matthews is Professor of Philosophy, Rutgers University, and member of the Rutgers Center for Cognitive Science. His research is focused in the philosophy of mind and in the foundations of cognitive science, specifically computational psycholinguistics. He has particular interests both in propositional attitudes, propositional attitude ascriptions, and their role in cognitive scientific theorizing and in formal learning-theoretic models of first language acquisition and their import for linguistic nativism. He has just completed a book-length manuscript, *The Measure of Mind*, which argues for a measurement-theoretic account of the attitudes.

James McGilvray teaches philosophy of mind and philosophy of language at McGill University in Montreal; he is also coordinator for McGill's cognitive science program. His work in recent years has been strongly influenced by Chomsky's effort to construct a naturalistic science of language. Books include *Chomsky: Language, Mind, and Politics* (1999) and *The Cambridge Companion to Chomsky* (as editor and contributor, 2005).

Jesse J. Prinz is Associate Professor in the Department of Philosophy at the University of North Carolina at Chapel Hill. He taught previously in the Philosophy-Neuroscience-Psychology Program at Washington University in St Louis. He is the author of *Furnishing the Mind: Concepts and their Perceptual Basis* (2002), *Gut Reactions: A Perceptual Theory of Emotion* (2004), and two other forthcoming books. He is also the editor of the forthcoming *Handbook to the Philosophy of Psychology*. His areas of research include concepts, emotion, moral psychology, the nature/nurture debate, and the neural basis of consciousness.

Geoffrey K. Pullum is Professor of Linguistics and Distinguished Professor of Humanities at the University of California, Santa Cruz. In addition to many books and articles in theoretical linguistics, his publications include a book of humorous and satirical essays about the study of language called *The Great Eskimo Vocabulary Hoax* (1991), and *The Cambridge Grammar of the English Language* (with Rodney Huddleston and others, 2002), a descriptive grammar which was awarded the Leonard Bloomfield Book Award from the Linguistic Society of America in 2004. He

was the Constance E. Smith Fellow at the Radcliffe Institute for Advanced Study at Harvard University in 2005–6, when his joint chapter with Barbara C. Scholz in this volume was completed.

Georges Rey completed his PhD in Philosophy at Harvard University in 1978. He works primarily on the foundations of cognitive science. He has published on issues of consciousness and qualia, concepts and intentionality, and the philosophy of linguistics. He is the author of *Contemporary Philosophy of Mind* (Blackwell, 1997), the editor (with Barry Loewer) of *Meaning in Mind: Fodor and His Critics* (Blackwell, 1991), and the section editor for cognitive science for the *Routledge Encyclopedia of Philosophy*. He has taught at SUNY Purchase, the University of Colorado, and has held visiting positions at MIT, CREA, the University of Split, the University of London, the Australian National University, and Stanford. He is presently Professor of Philosophy at the University of Maryland at College Park.

Richard Samuels is Lecturer in Philosophy at King's College, London. His research focuses primarily on issues in the philosophy of psychology and the foundations of cognitive science. He has published papers on nativism, cognitive architecture, evolutionary psychology and the implications of empirical psychology for our understanding of human rationality. He is currently completing a book on cognitive architecture.

Barbara C. Scholz lives in Santa Cruz, California, and held the Frieda L. Miller Fellowship at the Radcliffe Institute for Advanced Study at Harvard University during 2005–6, when her joint chapter with Geoffrey K. Pullum in this volume was completed. She publishes articles in journals of philosophy, psychology, linguistics, and psycholinguistics. Her particular interests lie in model-theoretic syntax and the philosophy of linguistic science.

John Tienson is Professor of Philosophy at the University of Memphis. He has published extensively on the foundations of cognitive science, including *Connectionism and the Philosophy of Psychology* (with Terence Horgan, 1996). He is currently working on a book in the philosophy of mind entitled *Phenomenal Intentionality*, with Terence Horgan and George Graham, Wake Forest University. He has recently published a dozen articles in the philosophy of mind related to the book, many with Horgan and Graham.

Christopher Viger has been assistant professor at the University of Western Ontario since 2002. He received his PhD in philosophy from McGill University in 1999 and has done postdoctoral work at Tufts University with Daniel Dennett; at the CUNY Graduate Center, with David Rosenthal; and at Rutgers University. His research areas are in philosophy of mind, philosophy of language, and cognitive science, with particular interest in the connection between language and thought in an attempt to find alternatives to the language of thought hypothesis. He has published in such journals as *Mind and Language*, *Synthese*, and *Philosophical Psychology*.

Ralph Wedgwood is a Lecturer in Philosophy at the University of Oxford and a Fellow of Merton College, Oxford. He has written several articles on metaethics and epistemology, and on related issues in metaphysics, philosophy of mind, and philosophy of language. He is currently finishing a book which gives a general theory of normative thought and discourse (that is, thought and discourse about what *ought* to be the case); this theory is based on the idea that "the intentional is normative" – that is, the idea that the nature of mental states that have "intentional content" can only be explained in partly normative terms.

Timothy Williamson is Wykeham Professor of Logic at Oxford University. He has also taught at Edinburgh University and Trinity College Dublin, and held visiting appointments at MIT, Princeton, ANU, and the University of Canterbury. He is the author of *Identity and Discrimination* (Blackwell, 1990), *Vagueness* (1994), and *Knowledge and its Limits* (2000), and of numerous articles in journals such as *Mind*, *Journal of Philosophy*, *Philosophical Review*, *Philosophy and Phenomenological Research*, *Mind and Language*, *Theory and Decision*, *Journal of Symbolic Logic*, *Journal of Philosophical Logic*, *Studia Logica*, and *Notre Dame Journal of Formal Logic*. He is a Fellow of the British Academy, the Royal Society of Edinburgh, and the Norwegian Academy of Science and Letters.

Preface

Robert J. Stainton

This volume is about debates in cognitive science. Yet it is part of a series called *Contemporary Debates in Philosophy*. How can it be both?

Let's begin with what cognitive science is. It is the interdisciplinary attempt to understand the mind, most especially the human mind. More specifically, one can think of cognitive science as having four branches: there are the behavioral and brain sciences, like psycholinguistics, neuroscience, and cognitive psychology; there are those social sciences that more or less directly inform us about the mind, like anthropology and sociolinguistics; there are formal disciplines like logic, computer science, and especially Artificial Intelligence; and, finally, there are parts of philosophy, especially philosophy of mind and language. The hallmark of cognitive science, in brief, is that it draws on the methods and results of all these branches, to attempt to give a global understanding of the mind.

To anticipate a worry, the idea obviously is not that philosophy is wholly contained in cognitive science. To pick only two examples, history of philosophy and political philosophy clearly aren't parts of cognitive science. What's more, even some parts of, say, philosophy of mind don't fit easily within cognitive science, e.g., issues about personal identity and life after death. The intersection, rather, is between certain sub-areas of philosophy and the other three branches.

From this definition alone we can immediately see why a debate can be both part of cognitive science and part of philosophy – for there is overlap between the two overarching fields. Some of the debates in this volume exemplify overlap of that kind. Brie Gertler and William Lycan debate about the nature and source of consciousness. Peter Carruthers, Jesse Prinz, and Richard Samuels debate the variety and extent of modular specialization in the human mind. Geoffrey Pullum and Barbara Scholz debate with Robert Matthews and James McGilvray about how language develops in the mind, and specifically about the role that an innate endowment plays. Such questions are core parts of traditional philosophy of language and mind, but they are

equally core parts of today's cognitive science. Thus there is an *intersection* of philosophy and cognitive science.

There is another way, however, in which a debate can be both philosophical and cognitive scientific. Many researchers accept that, though philosophy and empirical science are not the same thing, nevertheless the two are continuous. According to this view, called "naturalism," there is no sharp dividing line where philosophy stops and empirical science begins. This isn't merely the claim, just made, that a question can fall into both domains (e.g., the nature of space and time is among the oldest philosophical issues, but it is also pursued by experimental methods). The additional idea is that work which is straightforwardly empirical can bear on long-standing "properly philosophical" questions, and vice versa. Debates in this volume which exemplify empirical results informing philosophy include Kirk Ludwig and Chris Viger on the nature and function of perception. Recent research on how human thermoreceptors work, for instance, suggests that it is not their job to give an accurate representation of temperature to the agent. Instead, the job of a thermoreceptor is to bypass accuracy in favor of immediate, limb-saving reactions – like withdrawing a hand from something hot. This can seem to suggest that a very long tradition in philosophical thinking about perception misconceives the phenomenon from the get-go. (Ludwig firmly resists this inference.) Or again, what rationality is, is an extremely long-standing philosophical issue. The empirical research that Gerd Gigerenzer brings to bear in his debate with David Matheson again suggests that the philosophical tradition has deeply misunderstood rationality's fundamental nature. Going in the other direction, from philosophy to empirical science, Timothy Williamson urges that knowledge should be as central to the scientific understanding of the mind as it is to philosophical epistemology. Knowledge, insists Williamson, is a fundamental mental state with an importantly different behavioral profile than well-grounded belief. Thus the interdisciplinary attempt to understand the mind – cognitive science – cannot leave knowledge out. (This, in turn, means that cognitive science cannot ignore what is outside the mind, since a state counts as knowledge only insofar as it corresponds to what obtains "out there.") As another example, Ray Jackendoff and Georges Rey differ sharply about the implications for the science of language/mind of metaphysical worries about what "really exists" independently of the mind. These are all four of them *interactions* between philosophy and cognitive science.

So, how could a debate be both part of philosophy and part of cognitive science? In many ways, actually. There are various sorts of intersections and untold interactions. Speaking of "untold," let me end with this. Whether an investigation into x will yield evidence relevant to y cannot be known a priori: it depends upon whether x and y turn out to be linked in interesting ways. One just never knows for sure, then, which curious facts might turn out to be deeply evidentially relevant to a problem one is working on. To my mind, it is this aspect of the intersection and interaction between cognitive science and philosophy – never knowing where the next big lead may come from – that makes work in this area so challenging, but also so exciting.

Enjoy.

Contents

JUST HOW MODULAR IS THE MIND?

CHAPTER ONE

The Case for Massively Modular Models of Mind

Peter Carruthers

My charge in this chapter is to set out the positive case supporting massively modular models of the human mind.[1] Unfortunately, there is no generally accepted understanding of what a massively modular model of the mind *is*. So at least some of our discussion will have to be terminological. I shall begin by laying out the range of things that can be meant by "modularity." I shall then adopt a pair of strategies. One will be to distinguish some things that "modularity" definitely *can't* mean, if the thesis of massive modularity is to be even remotely plausible. The other will be to look at some of the arguments that have been offered in support of massive modularity, discussing what notion of "module" they might warrant. It will turn out that there is, indeed, a strong case in support of massively modular models of the mind on *one* reasonably natural understanding of "module." But what really matters in the end, of course, is the substantive question of what sorts of structure are adequate to account for the organization and operations of the human mind, not whether or not the components appealed to in that account get described as "modules." So the more interesting question before us is what the arguments that have been offered in support of massive modularity can succeed in showing us about those structures, whatever they get called.

1 Introduction: On Modularity

In the weakest sense, a module can just be something like a dissociable functional component. This is pretty much the everyday sense in which one can speak of buying a hi-fi system on a modular basis, for example. The hi-fi is modular if one can purchase the speakers independently of the tape-deck, say, or substitute one set of speakers for another for use with the same tape-deck. Moreover, it counts towards the modularity of the system if one doesn't have to buy a tape-deck at all – just

purchasing a CD player along with the rest – or if the tape-deck can be broken while the remainder of the system continues to operate normally.

Understood in this weak way, the thesis of massive mental modularity would claim that the mind consists entirely of distinct components, each of which has some specific job to do in the functioning of the whole. It would predict that the properties of many of these components could vary independently of the properties of the others. (This would be consistent with the hypothesis of "special intelligences" – see Gardner, 1983.) And the theory would predict that it is possible for some of these components to be damaged or absent altogether, while leaving the functioning of the remainder at least partially intact.

Would a thesis of *massive* mental modularity of this sort be either interesting or controversial? That would depend upon whether the thesis in question were just that the mind consists entirely (or almost entirely) of modular components, on the one hand; or whether it is that the mind consists of *a great many* modular components, on the other. Read in the first way, then nearly everyone is a massive modularist, given the weak sense of "module" that is in play. For everyone will allow that the mind does consist of distinct components; and everyone will allow that at least some of these components can be damaged without destroying the functionality of the whole. The simple facts of blindness and deafness are enough to establish these weak claims.

Read in the second way, however, the thesis of massive modularity would be by no means anodyne – although obviously it would admit of a range of different strengths, depending upon *how many* components the mind is thought to contain. Certainly it isn't the case that everyone believes that the mind is composed of a great many distinct functional components. For example, those who (like Fodor, 1983) picture the mind as a big general-purpose computer with a limited number of distinct input and output links to the world (vision, audition, etc.) don't believe this.

It is clear, then, that a thesis of massive (in the sense of "multiple") modularity is a controversial one, even when the term "module" is taken in its weakest sense. So those evolutionary psychologists who have defended the claim that the mind consists of a great many modular components (Tooby and Cosmides, 1992; Sperber, 1996; Pinker, 1997) are defending a thesis of considerable interest, even if "module" just *means* "component."

At the other end of the spectrum of notions of modularity, and in the strongest sense, a module would have all of the properties of what is sometimes called a "Fodor-module" (Fodor, 1983). That is, it would be a domain-specific innately-specified processing system, with its own proprietary transducers, and delivering "shallow" (nonconceptual) outputs (e.g., in the case of the visual system, delivering a $2^1/_2$-D sketch; Marr, 1983). In addition, a module in this sense would be mandatory in its operations, swift in its processing, isolated from and inaccessible to the rest of cognition, associated with particular neural structures, liable to specific and characteristic patterns of breakdown, and would develop according to a paced and distinctively-arranged sequence of growth.

Let me comment briefly on the various different elements of this account. According to Fodor (1983) modules are domain-specific processing systems of the mind. Like most others who have written about modularity since, he understands this to mean that a module will be restricted in the kinds of content that it can take as

input.[2] It is restricted to those contents that constitute its *domain*, indeed. So the visual system is restricted to visual inputs; the auditory system is restricted to auditory inputs; and so on. Furthermore, Fodor claims that each module should have its own transducers: the rods and cones of the retina for the visual system; the eardrum for the auditory system; and so forth.

According to Fodor, moreover, the outputs of a module are *shallow* in the sense of being nonconceptual. So modules generate *information* of various sorts, but they don't issue in *thoughts* or *beliefs*. On the contrary, belief-fixation is argued by Fodor to be the very archetype of a *non*modular (or holistic) process. Hence the visual module might deliver a representation of surfaces and edges in the perceived scene, say, but it wouldn't as such issue in *recognition* of the object as a chair, nor in the *belief* that a chair is present. This would require the cooperation of some other (nonmodular) system or systems.

Fodor-modules are supposed to be innate, in some sense of that term, and to be localized to specific structures in the brain (although these structures might not, themselves, be local ones, but could rather be distributed across a set of dispersed neural systems). Their growth and development would be under significant genetic control, therefore, and might be liable to distinctive patterns of breakdown, either genetic or developmental. And one would expect their growth to unfold according to a genetically guided developmental timetable, buffered against the vagaries of the environment and the individual's learning opportunities.

Fodor-modules are also supposed to be mandatory and swift in their processing. So their operations aren't under voluntary control (one can't turn them off), and they generate their outputs extremely quickly by comparison with other (nonmodular) systems. When we have our eyes open we can't help but see what is in front of us. And nor can our better judgment (e.g., about the equal lengths of the two lines in a Müller-Lyer illusion) override the operations of the visual system. Moreover, compare the speed with which vision is processed with the (much slower) speed of conscious decision making.

Finally, modules are supposed by Fodor to be both isolated from the remainder of cognition (i.e., encapsulated) and to have internal operations that are inaccessible elsewhere. These properties are often run together with each other (and also with domain specificity), but they are really quite distinct. To say that a processing system is *encapsulated* is to say that its internal operations can't draw on any information held outside of that system. (This isn't to say that the system can't access any stored information at all, of course, for it might have its own dedicated database that it consults during its operations.) In contrast, to say that a system is *inaccessible* is to say that other systems can have no access to its internal processing, but only to its outputs, or to the results of that processing.

Note that neither of these notions should be confused with that of *domain specificity*. The latter is about restrictions on the input to a system. To say that a system is domain specific is to say that it can only process inputs of a particular sort, concerning a certain kind subject-matter. Whereas to say that the processing of a system is encapsulated, on the one hand, or inaccessible, on the other, is to say something about the access-relations that obtain between the internal operations of that system and others. Hence one can easily envisage systems that might *lack* domain

specificity, for example (being capable of receiving any sort of content as input), but whose internal operations are nevertheless encapsulated and inaccessible (Carruthers, 2002a; Sperber, 2002).

2 What Massive Modularity Could Not Be

It is obvious that by "module" we can't possibly mean "Fodor-module," if a thesis of massive mental modularity is to be even remotely plausible. In particular, some of the items in Fodor's list will need to get struck out as soon as we move to endorse any sort of central-systems modularity, let alone entertain the idea of *massive* modularity.[3] If there are to be conceptual modules – modules dealing with commonsense physics, say, or commonsense biology, or with cheater-detection, to name but a few examples that have been proposed by cognitive scientists in recent decades – then it is obvious that modules cannot have their own proprietary transducers. Nor can they have shallow outputs. On the contrary, their outputs will be fully-conceptual thoughts or beliefs.

Domain specificity also needs to go, or to be reconceptualized in terms of functional rather than content domains, in the context of a thesis of massive modularity. Although it may well be the case that *many* modules are domain specific, it can't be the case that *all* are, if the thesis that the mind is built exclusively or almost exclusively out of modules is to be at all believable.[4] Consider practical reasoning, for example. This is plausibly a distinct system of the mind, with a significant innate component, whose internal operations might be encapsulated from and inaccessible to the remainder of cognition (Carruthers, 2004a). And it is a system whose basic architecture is probably very ancient indeed, being common even to insects as well as to ourselves and other mammals (Carruthers, 2004b). But it plainly can't be domain specific, since in order to do its job it will have to be capable of receiving any belief, and any desire, as input.

Swiftness of processing also needs to go, in the context of massive modularity, except perhaps in comparison with the speed of *conscious* thought processes, if the latter are realized in cycles of modular activity, as Carruthers (2002a) has maintained. For if the mind is *massively* modular, then we will lack any significant comparison-class. Fodor-modules were characterized as swift in relation to *central* processes; but a massive modularist will maintain that the latter are modular too. However, it looks as if the claim of mandatory operation can be retained. Each component system of the mind can be such that it automatically processes any input that it receives. And certainly it seems that some of the alleged central modules, at least, have such a property. As Segal (1998) points out, we cannot help but see the actions of an actor on the stage as displaying anger, or jealousy, or whatever; despite our knowledge that he is thinking and feeling none of the things that he appears to be. So the operations of our mind-reading faculty would appear to be mandatory.

What of claims of innateness, and of neural specificity? Certainly one *could* maintain that the mind consists almost exclusively of innately channeled processing systems, realized in specific neural structures. This would be a highly controversial claim, but it wouldn't be immediately absurd. Whether this is the *best* way to develop