

REVIEW SYMPOSIUM

MAKING SENSE OF EMPIRICISM?

Jesse J. Prinz, *Furnishing the Mind: Concepts and their Perceptual Basis*.
Cambridge, Mass.: MIT Press, 2002. Pp. x + 358. US\$40.00 HB.

Reviewed by Jonathan M. Weinberg

In his *Furnishing the Mind*, Jesse Prinz offers us a novel and scientifically-informed defence of concept empiricism. Prinz updates the old thesis that *nihil est in intellectu quod non fuerit in sensu* as a claim about what concepts are and where they come from: “all (human) concepts are copies of perceptual representations” (p. 108). Perception is the fundamental source of mental representation, and concepts are basically re-deployed copies of such perceptual representations.

Prinz’s writing is admirably clear and direct, and strives refreshingly at all times to get to the meat of whatever matter is under consideration. He offers his main arguments swiftly and with a minimum of superfluous window dressing. This directness can be a mixed blessing, however. The book weighs in at some three hundred pages of main text, and though that is not exactly a slender volume, Prinz has a *lot* of ground to cover in that space. As a result, anyone’s particular favourite patch of that ground may strike her as a bit thinly covered. In this review, I will set forth what I take to be the basic strategy of *Furnishing the Mind*, and suggest some places where I, at least, thought that a slower tour of the landscape might have been in order.

One can parse the overall argument of *Furnishing* into four unequal parts: first, an enumeration of the key desiderata of a theory of concepts (Chapter 1); second, the critical evaluation of extant theories of concepts by those desiderata (Chapters 2, 3 and 4); third, the articulation of Prinz’s preferred rival theory of ‘proxytypes’ (Chapters 6, 9, 10 and 11); and fourth, the defence of concept empiricism, especially from within the framework of proxytype theory (Chapters 5 and 7). Chapter 8 is a survey of attempts to resist nativist arguments, but since Prinz’s form of empiricism is not committed to any strong form of antinativism, the chapter is more a matter of pure defence – just in case someone else takes empiricism to have stronger commitments here. It is a useful synopsis of recent antinativist



thought, though, and anyone wishing to fashion herself a nativist today should know what her first lines of response are to the concerns raised there.

In the first stage of *Furnishing*, Prinz enumerates seven key desiderata for a theory of concepts: *scope*; *intentional content*; *cognitive content*; *acquisition*; *categorisation*; *compositionality*; and *publicity*. These desiderata all characterise fairly fundamental and widely agreed facts about human concepts. Briefly, we can understand the *scope* desideratum as the requirement that the theory accommodate the vast and varied range of our concepts. *Intentional content* and *cognitive content* are theories of reference and of sense. The *acquisition* desideratum is an account of how we get concepts in the first place. *Categorisation* is actually two related phenomena: ‘category identification’, our capacity to predicate concepts of appropriate particulars; and ‘category production’, our capacity to characterise members of categories, for instance, in terms of what other concepts they are likely to fall under. *Compositionality* and *publicity* are of course two of the more famous Fodorian constraints on a theory of concepts, concerning how we can generate new concepts almost without limit from ones already required, and how we generally share the same concepts.

So, how do the standard theories of concepts fare, when measured by these desiderata? Each seems designed to handle one set of explananda at the expense of others. Exemplar theories have a nice story about categorisation, for example, but fall down on compositionality (we all may have the concept ‘pet walrus’, but who’s ever experienced one?), and publicity (what two people will have the exact same set of dog exemplars)? Yet atomism, while stronger on compositionality and publicity, offers no account whatsoever of categorisation or acquisition. *Mutatis mutandis* for other candidate theories, such as the theory theory (great on concepts with essences but trouble otherwise) or such classics as imagism or definitions (which both have severe scope problems). Prinz’s hope is not to prop up one of these accounts and try to ignore their failings, but rather to offer a new, hybrid alternative that builds on the strengths of its forebears.

Prinz’s alternative, inspired by the research of L.W. Barsalou, starts by allowing in *all* the kinds of representations that previous theories have utilised, such as imagism’s mental pictures, prototype theory’s features, exemplar theory’s particular representations, and some of the more essentialist and discursive properties. For anything that we may be able to think about, we may have a vast amount of such representations associated with it. Just think of all the images of dogs you have seen or imagined, all the various associations you have with dogs, your tacit commitment to all dogs having a shared biological essence, any stories you’ve heard involving the

word ‘dog’ and so on – all of these doggy representations are candidates to play a role in the concept ‘dog’. But we surely do not use *all* of that information every time, or even any time, we think about dogs. Rather, at any time we wish to think about dogs, we assemble some subset of these representations to stand in for the whole of our understanding of the canine. The ‘proxytype’ we construct on a given occasion is context-sensitive, dependent on our cognitive needs of the moment. Prinz takes this context-sensitivity to give proxytype theory significant strength in explaining categorisation, for if “one is looking for dogs in the arctic tundra, one can call up a representation of a typical sled dog”, or if “one is looking for a guard dog, one can use a representation of a more ferocious breed” and “if one reads a news report about a dog that is five feet tall, one probably constructs a new representation on the fly . . .” (pp. 149–150). Supplement this account of what proxytypes are with the atomist’s informational semantics as an account of how they refer, and we have proxytype theory.

Much of the second half of the book is devoted to showing how proxytype theory can handle the seven basic desiderata. Chapter 11, for example, aims to meet Fodor’s compositionality arguments against prototypes and exemplars head-on, by showing how we can indeed construct appropriate new proxytypes for conceptual combinations to the extent that we actually do so. His account presupposes a relaxing of publicity that takes place in Chapter 6, such that it is sufficient if we overlap frequently in our default proxytypes, even if we do not always and everywhere share concepts. *Furnishing* has perhaps the best account on offer for how a non-atomist can approach this challenge.

Finally, if one could show that perceptual representations were the only representations we needed for a full array of proxytypes, then proxytype theory could form the basis for a new concept of empiricism. Much of the challenge here is in wrestling with the hard cases that have plagued empiricism since the early days of modern philosophy. Here the empirical lends aid to the empiricist, for Prinz avails himself of the robust notion of the perceptual that is common in today’s psychology, but which previous generations of empiricists would not have allowed themselves. Psychology motivates his theory, not epistemology, so he need not tie perception to sense-data or the like. Rather, he cashes out perception with the notion of dedicated input systems (p. 115). And, since science has indicated that we do have dedicated input systems that register causation, for example, proxytype theory will not have Hume’s difficulties in allowing for an empiricist concept of causation that is more than mere constant conjunction. Since proxytype theory can allow that any and all perceptual

representations can enter into a concept, and the robust contemporary view of perception includes the likes of causation, Prinz concludes that we have good grounds to be optimistic about concept empiricism.

But is that optimism warranted? First, I fear that Prinz misinterprets the rules of the game here, in a way that stacks the deck for proxytypes. Although a full theory of concepts should ideally explain all these desiderata, it does not follow that the theory should do so *only in terms of what sorts of things concepts are*. The ontology of concepts is surely a central part of one's full theory, but it need not *comprise* that full theory. If one's ontology of concepts must do all the heavy lifting, though, it is easy to see how proxytypes' everything-including-the-kitchen-sink nature will give them a leg up on the competition. But consider an analogy: a complete theory of automobiles would include an account of how cars propel themselves, what range of maneuverability they have and what determines how maneuverable a particular car will be, how cars are able to get into minor crashes with some degree of safety for their passengers, and so on. And explaining those aspects of automobiles will advert primarily to characteristics of cars themselves. But a complete theory of cars will need to include an account of other phenomena as well, such as what cars typically do when they approach red lights, or who tends to own which kinds of cars. Yet surely a theory of these aspects of automotive reality will have to relate cars to other parts of the world, such as traffic laws and financial institutions. So a theory of automobiles is better to the extent that it offers some account not just of car-intrinsic but also of these more extrinsic-but-car-involving phenomena. Similarly, a theory of concepts should at some point give an account of categorisation, or acquisition, or intentionality, but not necessarily with machinery built directly into the concepts themselves. It is simply not a strike against a theory of what concepts *are* if it does not also include a theory of some aspect of what concepts *do*. Now, if one can argue that a particular ontology of concepts would *preclude* any possibility of embedding that theory into a more complete theory of concepts, then we would indeed have a reason to reject that theory. But more often in *Furnishing* we are asked to weigh theories that are meant to be merely ontologies of concepts as if they were meant to be full theories. The atomists aren't offering an account of categorisation, for example, but there's no reason to think that they couldn't make use of all the machinery that the proxytype theorist does in offering such an account. They would just deny that all the feature lists and recorded exemplars, and so on, are actually a part of our concepts, as opposed to being associated with them.

Another worry one might have is whether there are still other desiderata that need to be listed as well, in addition to the seven discussed

above. Prinz does address the question of whether specifically linguistic desiderata should be included (he thinks that too many questions would be begged if they were), but we get no general discussion of how exhaustive this list is meant to be. It would be inappropriate to ask Prinz for any sort of overarching philosophical argument that these seven and only these seven desiderata are applicable to a theory of concepts, because Prinz is admirably naturalistic in his approach to these questions – the theory must be responsive to the explananda, and what explananda there are depends on the contingent ways we happen to discover the world to be. Nonetheless, it may seem that some important empirical desiderata may have been left out. In particular, although Prinz is happy to talk about concepts as ‘thought constituents’ and that much of the import of a theory of concepts derives from its potential relevance to a theory of thinking and cognition in general, he proposes no desideratum for this *thought-friendliness* of concepts. Yet I fear that we have some reason to suspect that proxytypes, as an ontology of concepts, will make it very hard to account for thought-friendliness.

The problem arises from the very context-sensitivity that Prinz takes to be a selling point of proxytypes. If I start to think about dogs, then proxytype theory says that I construct some context-appropriate proxytype of dogs to use for my ruminations. But suppose I reach some final conclusion of my thought-process, perhaps the conclusion that Prinz likes dogs. (He seems to use them a lot in his examples, after all.) How is that conclusion to be stored, such that it represents a belief concerning *all dogs*, and not merely a belief concerning *dogs-considered-under-such-and-such-a-proxytype*? This is not a question about intentionality – let us grant that each dog-proxytype refers to all and only dogs – but about the nature of reasoning. Prinz’s theory includes some machinery, the “long-term-memory network” (see pp. 144–148), that appears meant to do some of this work. These networks are networks between perceptual representations, with prototype features linked to representations of exemplars linked to linguistic representations, and so on. Suppose that I am using a simple single visual model of a particular dog as my dog-proxytype of the moment (Prinz states that this is possible, p. 149). The theory needs a way to distinguish between that particular visual model serving as a dog-proxytype, or as a beagle-proxytype, or as a Snoopy-proxytype, or even serving (as it is in this discussion) as a proxytype-proxytype. Again, these distinctions should reflect not only different mind–world relations, but also different ways in which the proxytypes can enter into our cognition even considered solipsistically. Perhaps they should be considered part of the cognitive content of the concepts.

We can see the same problem again when we consider formal inferences. As Fodor and Pylyshyn have argued, mental representations need to support inferential coherence. We are able to reason from p and *if p then q* to q , but that inference only makes sense if it is the same p 's and q 's throughout. But if the p -representation in the first premise is distinct from the p -representation in the second, then instead of instantiating *modus ponens* we would be instantiating a non-sequitur. It's hard to make a piece of reasoning out to be an inference, if we don't have a way of saying what is shared across different pieces of information. In addition to the myriad individual perceptual representations we might have about dogs, and the big interlocking soup of all that information, we still need something else to hold it in place, to serve as a basis for the reasoning that calls upon and adds to that network. What I am inclined to say is missing, then, is the concept itself.

Prinz shows a partial recognition of this issue at one point, writing that "we should hope that the mind is furnished with rules that combine and manipulate thoughts in a way that can, in some situations, be fairly reliable in preserving truth"; and such rules require that "thoughts must have something like logical form" (p. 181). But it is unclear that we really can get by with something only *like* logical form, instead of the real thing. Prinz sketches how we might do something like negation, suggesting that we can form the proxytype for *not-A* by, in essence, flipping all the values on a proxytype for *A*. But that is not so much *not-A* as *un-A*, and does not have the structure necessary to support, say, disjunctive syllogism. His discussion of logic wanders off at that point, and I fear that it conflates *having* logical form, with being something that can be treated *as if* it has logical form. He suggests that we may use "mental analogues of Venn diagrams" (p. 184) to reason with quantifiers. But a picture of a Venn diagram has no more logical structure to it than ' $(x)Fx$ ', considered as an uninterpreted string. If we can reason with such diagrams or such strings, it is because we can mentally assign them the logical form they need. But with a theory of concepts, we would want exactly such an account of such a capacity to assign logical form to fall out of an explanation of thought-friendliness – an explanation which we are not offered. We should not be surprised if proxytypes turn out to offer little help here, because they too were designed to handle one set of phenomena while ignoring others. Prinz's overall interest is in concepts as detectors – as the sorts of mechanisms that enable us to take in part of the world and determine what sort of thing it is. But detection is only part of the story of concepts, albeit an absolutely central part. What we do with concepts after the detection has taken place should set a significant part of our research agenda as well.

Finally, Prinz is only partially successful in his defence of conceptual empiricism against problem cases of non-perceptual concepts. His Chapter 7 is an all-too-quick tour of various moves that an empiricist could make, but other than the section on causation, none of them are fleshed out enough to be at all convincing to anyone not already inclined towards empiricism. His account of the concept *electron* depending on the concept *trace photograph* (i.e. of a cloud chamber) sounds like some very old-fashioned philosophy of science, and certainly is unconvincing psychology. (How did the scientists think about electrons before the invention of the cloud chamber, so that they could figure out that a cloud chamber was a good device for looking for electrons?) And the concept of necessity, surely one of the great bulwark concepts of rationalism, is startlingly absent from his discussion. But perhaps a proxytype theory that was revised to accommodate the worries about thought-friendliness, above, would have the resources to address these problem cases more thoroughly. Nonetheless, until such a revised proxytype theory comes along, with its accordingly expanded catalog, I would advise shoppers not to do all their furnishing from the empiricist's shop.

Department of Philosophy
Indiana University
Bloomington, IN 47405
USA

Reviewed by Daniel Yarlett and Michael Ramscar

In this ambitious book, Prinz sets himself the task of giving a comprehensive account of the nature of our concepts consistent with the basic tenets of empiricism – the idea that our conceptual knowledge is fundamentally derived from and grounded in our perception of the world. Although empiricism has a long history in philosophy, and in psychology too, albeit such that psychologists tend to refrain from explicitly endorsing such doctrines, Prinz paints the current intellectual climate in both disciplines as being largely inimical to the idea. Prinz lays out his theory against a landscape that is depicted as being largely dominated by nativist theories which hold that the interesting aspects of human knowledge are biologically inherited, instead of acquired through perceptual experience. According to Prinz, “most cognitive scientists think that strong forms of nativism are not only tenable, but also ineluctable” (p. 110). Thus the task,

as Prinz sees it, is to rehabilitate empiricism so that it may be admitted into the modern scientific community.

Just how accurate is this picture? While the writings of some prominent authors in the field may incline one to the assessment that nativism dominates – the names Chomsky, Fodor, Pinker and Carey spring most obviously to mind – this view is far less universally endorsed than Prinz suggests, especially amongst empirical researchers. Although there is a widespread support for nativist theses in the fields of linguistics and, curiously, developmental psychology, the majority of those psychologists who focus on perception and cognition assent to the idea that much of our knowledge is acquired via experience and mediated by general learning mechanisms. Indeed, the researches of a great number of animal and human learning theorists have, over the last century or so, provided powerful demonstrations of the degree to which associative learning processes offer satisfying explanations of (and testable predictions for) a range of psychological phenomena; these findings form the bedrock of most research into perception, memory and learning. Painting the cognitive sciences as inherently nativist is a forgivable rhetorical conceit on Prinz's part – who wouldn't prefer to see the theoretical edifice that one intends to climb perceived as more of a mountain than a molehill? But we feel it worth noting at the outset that the picture of mainstream cognitive psychology he paints is questionable (paradoxical even, given that much of the evidence Prinz draws from the literature of psychology and the neurosciences advanced in support of his theory is empirical – and empiricist – in nature).

What of the theory? Prinz terms the specific doctrine that he wishes to defend *concept empiricism*, which he defines as the thesis that “all (human) concepts are copies or combinations of copies of perceptual representations” (p. 108). This definition is thus an updated version of traditional Imagist theses, which propose that thoughts correspond to mental images. Forms of this doctrine have variously been advocated by Aristotle, Locke, Hume, and, more recently, Russell and Ayer. In order to distinguish Prinz's thesis from more traditional formulations, a lot hangs on the exact nature of ‘perceptual representations’. In what way are these distinct from Images or Sensations? In answering this question Prinz has recourse to the notion of a *sensory system*; he defines a perceptual representation as any representation derived from one of the senses. This in turn begs the question of what is a ‘sense’. According to Prinz there are two main ways of defining the nature of our senses that have previously been proposed, although neither gets it quite right. The first potential definition is that something counts as a sense because it is a modular system in the Fodorian manner. Under this approach the senses are construed as independently functioning

subsystems of the brain that operate quickly, are domain- or input-specific, are associated with specific neural architectures, and are ‘informationally encapsulated’ (meaning that their computations occur independently of the information available in other systems of the brain). The second potential definition is that the senses are fundamentally receptive in nature, in that they respond passively to environmental inputs. This is intended to contrast the senses with the intellect, which can be spontaneously engaged in the pursuit of a specific goal.

Prinz’s account of the senses, and hence of perceptual representations, is a marriage of these two approaches. According to him, the modularity proposal falls down primarily because of empirical evidence that perception is far from informationally encapsulated. There are many experiments showing that perceptions in one modality can affect perception in another modality, such as in colour, phoneme, object and word identification tasks. Recent studies (by Boroditsky in particular) have also shown that language can have strong effects on both perception and memory, further undermining the claim that perception is constituted of fully-encapsulated systems. Prinz also argues that the receptivity definition is insufficient on its own, both because the intellect is not always spontaneously engaged (certain thoughts occur irrespective of whether we intend them to), and because the senses can be spontaneously recruited in order to acquire specific information (as in the case of visual search). Prinz contends, however, that there is an element of truth in both accounts and thus marries them by claiming that “the senses are *dedicated input systems*” (p. 115). This definition decomposes into three more or less independent claims. The claim that the senses are systems corresponds to the assertion that they have independent neural pathways and are thus functionally discrete (this requirement seems to fall short of stipulating complete encapsulation, however). The claim that the senses are input systems means that their prime purpose is to respond to variation in the state of the environment and make information about this variation available to the brain. And the idea that sensory systems are dedicated means that each responds to its own proprietary class of inputs, and represents this input in a unique fashion.

Prinz’s characterisation of the senses is an intriguing one and is, for the most part, convincingly argued for in a brisk and robust fashion. At a minimum his treatment does a good job of exploiting research in the cognitive sciences in order to add more meat to the bones of traditional Imagist theories, and it would seem likely to provoke further debate in the future. Perhaps the most controversial part of Prinz’s characterisation is the claim that the senses each use a different form of representation. This claim is very much a speculative one given the limited nature of our

current understanding of mental representation, and only future empirical work will be able to convincingly militate between 'multimedia' accounts of mental representation of the sort favoured by Prinz, and the 'common-code' or amodal representations assumed by many rationalists and, until recently, presupposed by the many cognitive scientists sold on the brain-as-computer metaphor.

Having outlined his theory of perceptual representations, Prinz moves on to describe the way in which these can be cognitively structured in order to give rise to concepts. A key notion he uses is that of 'long-term memory networks', which consist of perceptual representations that are linked to one another in a variety of ways, according to different 'linking principles.' These principles can be of a range of different types – including transformation, binding, situational, hierarchical, and predicative links – and are individuated by the functional role that they play in cognitive processing.

There are two main shortcomings in Prinz's discussion of the nature of the linking principles. The first concerns the way in which specific types of links are supposed to be derived from perceptual experiences. Although Prinz gives some examples of the way that specific links might be acquired – for example, he discusses the case of representations of a gorilla standing still and beating its chest becoming associated via a *transformation* link, because the same object is perceived as changing its form – these are not always as detailed as one may have hoped. In particular, the notion of a predicative link plays an important role in Prinz's theory, yet its precise nature is described in a somewhat elliptical fashion. Predicative links are intended to capture the idea that we are sometimes disposed to transfer information that applies to one perception to another; for example, when the knowledge that Fido and Rover enjoy digging for bones leads us to believe that Spot, a canine of recent acquaintance, will also enjoy passing the days in a similar fashion. According to Prinz, such representations become "stored together because they are quite similar"; this is the "matching principle", and how they are connected "can be called a *predicative link*" (p. 145). Unfortunately, relying on similarity to do the work of induction is a notoriously slippery slope, and given the notion's importance to his theory, it is unfortunate that Prinz did not spend more time dealing with this issue. Further work in this direction would constitute an important addition to his theoretical achievement.

The second shortcoming of Prinz's discussion of the linking principles is that it is unclear whether the link-types discussed are intended to be an exhaustive list, or whether new link-types can be postulated as circumstance demands. If the latter is true, this raises concerns about the

falsifiability of Prinz's theory: assuming the basic correctness of concept empiricism, what could count as evidence against the existence of long-term memory networks? Overall it would have been useful to see a more explicit statement about the link-types presupposed by the theory, and the exact manner in which they are to be individuated.

Along with long-term memory networks and linking principles, the other central notion to Prinz's theory is that of a 'proxytype'. Proxytypes are "mental representations of categories that are or *can be* activated in working memory" (p. 149). The idea is that only specific subsets of an individual's long-term memory store – in the form of proxytypes – can be recruited on a given occasion, and that the recruited structure is what corresponds to the concept active in that instance. Proxytypes are intended to serve two main purposes: first to recognise or classify items in the world (they thus function as detection and tracking mechanisms); and second to form parts of thoughts when activated. This latter point implies that thinking is a form of perceptual simulation, which is an intriguing idea that has recently been championed by Larry Barsalou (whose influence Prinz acknowledges).

As previously discussed, Prinz's theory is promiscuous in its postulation of links between representational items in memory. A serious challenge faced by any theory of this sort is that of concept individuation, and this is perhaps where the theory is on shakiest ground. According to Prinz, virtually "any concise subset from such a network can do the job" of serving as a proxytype (p. 149). Given this, unless a way of successfully carving up these sets of representational items into discrete chunks of appropriate size and function can be found, then the theory will ultimately fail to successfully individuate 'concepts', which is one of the desiderata Prinz sets out at the beginning of his book.

This problem is further highlighted when one considers the way in which context can affect the way in which conceptual knowledge is recruited. Prinz explicitly acknowledges that context strongly interacts with the representations that are activated in a given instance. "Context determines what proxytype is used in working memory on any given occasion" (p. 149). While it is good to see a theory of conceptual structure attempt to address the problem of context-dependency – this is a problem with which many psychologists struggle – it does raise problems for the theory in terms of the individuation of concepts, and their putative publicity (the idea that concepts must be shared because we can use them to communicate).

Prinz's answer to these problems is the notion of a 'default proxytype', which is intended to form our core representation of natural kinds in the

world. Default proxytypes are meant to capture the common instantiations of features that we observe, and accordingly, what “ultimately determines whether a feature is included in a default proxytype is how frequently one represents a category as having that feature . . . [and] the frequency with which a feature is used can be affected by its cue validity, category validity, perceptual salience, or conformity to theories”; all “exert some influence” (p. 155). What Prinz fails to convince us of is that this doesn’t result in our having a different concept every time we have a thought. In the absence of a detailed description of the linking principles that can relate perceptual representations, and the ways in which these are derived from experience, there are no convincing reasons to believe that either standard or default proxytypes can be properly individuated. (Given so many links in memory, why aren’t a few more representations recruited on one occasion, and a few less on another?) This is a serious problem for the theory Prinz presents, and it seems soluble in only one of two ways: either by providing a more precise definition of how it is that specific link-types can be used to circumscribe individual proxytypes (which seems to be a non-trivial issue), or else by embracing the more radical position that talking about ‘concepts’ is simply a rhetorical conceit for conveniently discussing human (and animal) knowledge, and thus that there may therefore be no strict identity that obtains between the knowledge recruited on any two occasions.

Prinz begins his book with a set of desiderata for concepts that spell out what a theory of concepts should give us. One way of regarding these desiderata is as an idealised description of the way in which a cognitive system might work (had we evolved or been designed in that way). While we appreciate Prinz’s ingenious efforts to reconcile the data in the psychological literature with these desiderata, our suspicion is that the fact that the desiderata and the data are so hard to reconcile could owe much to the fact that, ultimately, the desiderata – and the delineable, determinable concepts that they presuppose – fail to characterise human concepts accurately. This is unlikely to prove a popular conclusion: first, given the near-consensus amongst philosophers that the ‘concept problem’ is well-posed; and second, given the particular challenges it raises with regard to explaining how it is that we do successfully communicate and understand one another without strict identity between our ‘concepts’ (that is, without ‘concepts’ in any real sense). However, in spite of the new challenges it raises, we believe that Prinz’s characterisation of conceptual knowledge is on the right track, and that the difficulties his theory has in meeting the desiderata are more a reflection on the shortcomings of the view of conceptual knowledge embodied in the desiderata than they are on the

theory itself. Human cognitive processes are notoriously probabilistic, non-optimal, vague, and downright undercooked. Prinz's desiderata lay out a cognitive architecture for the mind as philosophers would have it. The truth may be that in attempting to live up to such standards, we are forced to furnish the mind with more than is strictly necessary to give an accurate account of human thought.

*Psychology Department
Stanford University
Jordan Hall
Stanford, CA 94305
USA*

Reviewed by Dan Ryder

In this wide-ranging book, Jesse Prinz attempts to resuscitate a strand of empiricism continuous with the classical thesis that all ideas are imagistic. His name for this strand is "concept empiricism" and he formulates it as follows: "all (human) concepts are copies or combinations of copies of perceptual representations" (p. 108). In the process of defending concept empiricism, Prinz is careful not to commit himself to a number of other theses commonly associated with empiricism, more broadly construed. For example, he is prepared to accept that there are innate concepts and/or knowledge, denies that what a concept means consists in the experiences that prompt us to use or create it, implies that cognitive architecture is not associationist, and offers no opinion on whether all knowledge claims must be justified by sensory experience. Those who await a full resurrection will have to wait a little longer – but in the meantime, Prinz's reconstructive surgery will tide you over. Although it falls short of miraculous, it is still impressive.

Prinz has brought a vast knowledge of the literature to bear on his project, from philosophy, psychology, and neuroscience. In fact, this book would serve as an excellent entrée for the philosopher into the scientific aspects of concept research, or for the scientist into philosophical concerns. Prinz writes with exemplary clarity, and wields his theory with aplomb in answering the many objections that have been raised against imagism. To take just one example, anyone who doubts that imagism can accommodate the large scope of human concepts would be well advised to read Chapter 7, which contains a wealth of ingenious suggestions for how imagism might handle difficult cases, including lofty concepts

such as *cause* and *truth*. His discussions of nativism (Chapter 8) and compositionality are also particularly illuminating.

The central theoretical construct in Prinz's theory of concepts is the 'proxytype', a group of imagistic/perceptual representations. Proxytypes have two main functions. First, they serve as detection and tracking mechanisms. Second, they can be activated in working memory to form part of a thought. According to Prinz, who cites Larry Barsalou as his main influence (see e.g. "Perceptual Symbol Systems", *Behavioral and Brain Sciences*, 1999), thinking is simulation of perception. To think about something, you put yourself into an imaginative state that resembles your perceiving it. In this way, proxytypes stand in for worldly items on an internal stage used for planning and prediction.

Proxytypes have parts (ultimately primitive perceptual/imagistic representations), and they themselves often form parts of larger imagistic or perceptual scenarios. This raises the question of how to individuate proxytypes; one cannot find a concept in any arbitrary set of perceptual representations, just as one cannot find an object in any arbitrary set of tiles in a mosaic. Prinz does not give us enough help here (with consequences I will worry about later), but the idea seems to be that legitimate groupings are determined by "links" in long-term memory. Most importantly, these links serve to 'call up' the rest of a proxytype when only part of it is tokened, e.g. an auditory image of barking might call up a visual image of a furry dog-shape via an intermodal "binding" link. Together they may form a *dog* proxytype. Links come in various types (e.g. binding, situational, transformational) and are distinguished by their functional roles. Prinz is very permissive about the internal psychological structure that a proxytype can have, and thinks that all the main types of psychological theories of concepts – prototype, exemplar, and theory theory – have insights to contribute on this matter.

Prinz offers a novel causal/informational theory of conceptual content (a 'psychosemantics'), that yields *two* types of contents for each complex proxytype – a real or intentional content, and a nominal or cognitive content. For example, the proxytype consisting of an auditory image of barking and a visual image of a furry dog-shape might have as its nominal content a certain doggyish appearance, and as its real content the biological kind *dog*. There is a one-to-one mapping of proxytypes onto nominal contents or appearances, but a many-to-one mapping of proxytypes onto real contents. Each of many ways one has of recognising a dog may constitute its own proxytype, with its nominal content being a particular dog appearance. But the real content of each of them is *dog*.

Prinz takes his main competitors to be definitionism, prototype theory, exemplar theory, theory theory, and informational atomism (having swiftly dismissed theories that deny concepts are mental particulars, *inter alia* those espoused by Peacocke, Dennett, Wittgenstein, and Brandom – rather too swiftly, perhaps). He aspires to no knockout punches but rather claims to win on points, as determined by how a theory stacks up against a list of desiderata set out in Chapter 1. Other things being equal, a theory of concepts ought to explain the scope of human concepts, their intentional contents (analogous to reference), their cognitive contents (analogous to sense), and both the compositionality and the publicity of these contents. It also ought to explain how concept acquisition is possible, and the role concepts play in categorisation (construed very broadly to include judgments of similarity and even inferences involving categories).

Although some will be inclined to quibble over this list of desiderata, I am willing to accept them. My complaint is that like an Olympic judge, Prinz has been over-generous to his own country's skater. A central flaw mars his theory's performance with respect to two fundamental desiderata: the explanation of categorisation (which Prinz takes to unify the theory), and the theory of content. The flaw lies in his account of concept individuation. First, consider the act of recognising or identifying someone. As Ruth Millikan has persuasively argued in *On Clear and Confused Ideas* (Cambridge, 2000), this act is not simply to react perceptually in the same way as on a previous encounter with that person, nor is it merely to wield the same representations again. The amnesiac, who is unable to form any new memories, may react perceptually in exactly the same way, and deploy exactly the same perceptual representations every time he sees you. But he does not thereby recognise or identify you. (Every time he says, 'hello, how very nice to meet you'.) Something further is required. Millikan suggests that the sign of someone recognising you is not their reacting in the *same* way as before, but rather their acting *differently*: by applying what they learned last time. They might, for instance, call you by your name. Without such an ability to take account of information gathered on previous encounters, as occurs in the amnesiac's case, there can be no recognition. The lesson that mere repetition is insufficient for recognition extends to concepts of kinds, and of anything else that one might re-identify as being the same again – this is water (again), this is a book (again).

Prinz understands categorisation in these terms (like Millikan, he speaks of 'tracking' individuals and kinds), so he owes us an explanation of the ability to re-identify. (Millikan herself takes a very different line, saying that a concept *just is* such an ability.) In order for a repre-

sentation, simple or complex, to explain this ability, it must possess some feature, a “sameness marker” (Millikan, p. 146) that serves as a signal to the cognitive system that “here is the same thing again”, so the cognitive system can attach new knowledge to that representation, to be carried forward to later encounters. Where proxytype theory falls down is in its account of sameness marking.

The atomist, for example, has an easily available sameness marker: either identity of vehicle (e.g. the same population of neurons) or identity of syntactic type (in the language of thought). It seems that neither of these is available to Prinz, in most cases. Consider, in proxytype theory, concepts individuated by their real contents, which enable us to track or re-identify real kinds (indeed, Prinz says this is the *point* of having proxytypes; p. 281). Such concepts come in the form of *multiple* complex images “linked” in long-term memory. On one occasion, I may wield a proxytype whose nominal content is a Doberman appearance, on another occasion I may wield a proxytype whose nominal content is a Chihuahua appearance, while on yet another occasion I may wield a proxytype “constructed on the fly” in an encounter with some sort of dog I’ve never seen before (p. 150). There is no vehicular or syntactic identity amongst these various tokens of the *dog* concept. Yet Prinz maintains that in wielding these very different proxytypes, I can (re-) identify three different animals as being the same, as *dogs*.

How do I accomplish this re-identification, despite the fact that I have “countless dog concepts”, i.e. individual proxytypes that denote dogs (p. 152)? What I do not have in my conceptual repertoire, according to the concept empiricist, is some abstract representation of dogness I deploy simultaneously, whose vehicular or syntactic identity could serve as a sameness marker. On Prinz’s account, as far as I can tell, a different identity marks sameness: the identity of the long-term memory network that gives rise to the proxytype tokened in a particular thought. (See p. 151, where Prinz uses this device to individuate the hunting, fatness, and gnu proxytypes in a complex image of hunting a fat gnu; see also pp. 161–162.) This account of sameness marking requires that long-term memory networks be individuable at a rather fine grain, which is highly problematic at best.

Suppose I activate a proxytype whose nominal content is a Doberman appearance, but whose real content is *dog*. This ‘Doberman-appearance proxytype’ is linked in long-term memory to a *lot* of things besides other kinds of dogs – my neighbours in 1978, blackness and brownness, the letter D, hypothyroidism, and warlocks (after Borong the Warlock, a famous Florida Doberman) – and thereby to many, many other proxytypes. Given this massive overlap of long-term memory networks, what

makes it the case that my Doberman-appearance proxytype arises from my dog network (and thus has the real content *dog*), rather than my hypothyroidism network, or my blackness network? After all, presumably it is usually only a *part* of a network that actively produces a token proxytype. (For example, it would be rather implausible to maintain that my stored Chihuahua-appearance proxytype played a causal role in producing my currently active Doberman-appearance proxytype!)

Prinz's answer seems to lie in the various *kinds* of links in long-term memory, where these kinds differ in their *functional roles* (pp. 145–148). Some links are, as it were, *intranetwork* (my term), e.g. what Prinz calls the 'predicative' link between my stored Doberman-appearance proxytype and my stored Chihuahua-appearance proxytype. Intranetwork links serve to individuate the dog network. Other links are *internetwork*, e.g. the 'situational' link between my Doberman-appearance network and my neighbour network. Take as an example of an intranetwork link the predicative link. According to Prinz, a predicative link endows me with a disposition to 'transfer' features from one representation to the other. First, I am puzzled about what perceptual features I am disposed to transfer between Doberman appearances and Chihuahua appearances. Not size, nor colour, nor most other things. Having two ears? But if that counts, then surely I am equally disposed to transfer that feature to a Siamese cat-appearance proxytype, so it fails to help isolate the dog network. What sort of transfer *could* isolate it? (Dogness would do it!) Second, surely the two proxytypes under consideration are also part of *distinct* networks, namely my Doberman network and my Chihuahua network. It seems the predicative link is both a mark of sameness and not a mark of sameness. Prinz's discussion of these crucial long-term memory links left me feeling profoundly unilluminated. As it stands, proxytype theory fails to account for sameness marking, and therefore fails to explain categorisation. *Contra* Prinz, atomism does better: though Prinz faults atomism for failing to explain categorisation because structureless concepts cannot explain discrimination, atomism just explains discrimination by appeal to structures *external* to concepts, e.g. beliefs.

Prinz *does* accept atomism for the primitive perceptual representations out of which proxytypes are constructed. For these representations, he can avail himself of the easy kind of sameness marking. For example, we have the edge-detectors on p. 274, where Prinz seems to rely on neural identity as a sameness marker. In a move that might be thought of as cheating, he reduces the impact of the network individuation problem by expanding his primitive base beyond what you might expect for an empiricist. For instance, he is prepared to include among perceptual primitives

bi-modal representations (p. 137), representations of Biederman's 'view-independent geons' (p. 140), and representations of objective size (p. 187). One begins to wonder if Prinz's empiricism is a bit tepid (his definition of a perceptual representation as "a representation in a dedicated input system", p. 115, would also seem to cast a very wide net).

The network individuation problem ramifies beyond Prinz's unifying desideratum of explaining categorisation. Another important strand in the book that suffers is his account of the real/intentional content of concepts. In keeping with his account of concepts as detection mechanisms, one essential element of Prinz's psychosemantics is causal covariance. For example, let us take a concept denoting Monarch butterfly. (The concept will also co-vary with the superficially identical Viceroy, but Prinz maintains that it does not denote *Viceroy* because it was not originally created in order to track Viceroy. Among the things it causally co-varies with, a concept represents its "incipient cause", the thing that caused its creation.) *Monarch* is a real rather than a nominal content. As we saw, this means that the Monarch concept is actually constituted by a number of *different* proxytypes linked in long-term memory (Monarch with wings closed, Monarch with wings open, etc.). Any *particular* tokening of the Monarch concept will involve only a small subset of these proxytypes, each of which causally co-varies only with (context dependent) contingent *signs* of the kind (p. 169). What really causally co-varies with Monarchs, then, are the proxytype members of the Monarch network, *taken together*. Obviously this requires that the Monarch network be individuable, an unsolved problem for Prinz, as we saw previously.

The problem is worsened, though. We can now see that this individuation must be accomplished *independently of the network's causal co-variation properties*, on pain of trivialising the psychosemantics. It turns out that individuation cannot depend upon the network's particular constituting proxytypes either, for those can vary (e.g. through learning, pp. 252–253), even very radically: "my *whisky* concept is constituted by different proxytypes in [possible] worlds with blue whisky" (p. 280). It seems to me that Prinz has dug himself a very deep hole here. His whole project seems to depend crucially upon being able to individuate long-term memory networks, yet he cannot individuate them by their causal relations to the world (and thus not by their real content), nor by their nominal content, nor by their constitutive atomic representations, and he is hard pressed to do so by their vehicles. What else is left? And remember, the individuation has to be sufficiently robust to do the job of sameness marking – the cognitive system has to be able to *interpret* this individuation

in order to organise newly acquired information, or Prinz's concepts cannot explain categorisation.

I certainly hope that Prinz can relieve his theory of the network individuation problem, either by solving it, or justifiably putting it aside. A repair to that part of the foundation would leave an impressive edifice indeed, with a richness of detail impossible to convey in a short review.

Philosophy and Cognitive Science
Indiana University
Sycamore 026, 1033 East 3rd Street
Bloomington, IN 47405
USA

Author's Response

By Jesse J. Prinz

In *Furnishing the Mind*, I argue that concepts are proxytypes, or perceptually derived and highly variable constructions in working memory, derived from long-term memory networks that serve as detectors for instances of categories. My defense of proxytype theory is incomplete. Some of the core claims require further evidence, and some of the theoretical constructs require further elaboration. Until both deficits are remedied, the safest attitude is one of caution. I am greatly indebted to the reviewers for exposing points that need special attention. Their commentaries are very generous, but they all politely advance serious concerns. If these concerns cannot be addressed, the theory I favour will not get out of the starting gate. Unfortunately, I can do no more than gesture at responses here. The concerns warrant much more attention in the future. My theory will either emerge better for these efforts, or it won't emerge at all.

Interestingly, the claim that concepts are perceptually derived gets relatively little discussion in these commentaries. I would have expected that empiricist thesis to be the most controversial part of proxytype theory. It turns out there are more pressing worries. I cannot possibly address all of the objections here, so I will focus on two. The first concerns concept individuation, and the second concerns the desiderata that I say any theory of concepts should seek to satisfy.

CONCEPT INDIVIDUATION

In the book, I follow Larry Barsalou in claiming that the same category is represented differently on different occasions. The category of dogs may be represented by a structured image of a beagle one time and a pug on another. How can these two representations be manifestations of the same concept? This is the crux of the concept individuation problem. I contend that the problem is solved by links connecting active representations to memory networks. In particular there are predicative links, which can be characterised by their functional roles. If a representation is predicatively linked to a memory network, features in the network are likely to be transferred to that representation. Two representations are manifestations of the same concept if they are predicatively linked to the same memory network.

Yarlett and Ramscar raise a worry about this. They recognise that predicative links are the conduit for induction, and induction is a hard nut to crack. There is a suggestion in the book that similarity plays a role in establishing predicative links. This, Yarlett and Ramscar rightly point out, cannot be the whole story. Similarity is not a perfect guide to induction. But I did not intend to imply anything that strong. Similarity plays the following role. When we encounter a new object (an unfamiliar beagle, for example), we search our memory for a similar item. The closest match may determine the category. Once a match is found, a representation of the perceived beagle will be predicatively linked to a long-term memory file. So similarity can license induction. But I do not think similarity is necessary or sufficient for forming predicative links. In some cases, we link a representation to a network with a central tendency that bears little resemblance to it (e.g. Frank Keil's raccoon, painted to look like a skunk). I do not offer a general theory of induction, or a theory of all the conditions under which we form induction licensing links. I wish I had a story. But until I do, I hope that my appeals to similarity will be regarded as an uncontroversial piece of that story.

Yarlett and Ramscar have another objection waiting in the wings. If categories are represented differently on different occasions, then there is little point to saying these representations are instances of the same concept. Concepts are usually presumed to be discrete and stable entities. That assumption is essential to explanations of communication. If concepts vary too much, they cannot be shared. In the book, I suggest that there are default representations of categories, which remain relatively stable and converge across individuals. Yarlett and Ramscar complain that I cannot help myself to such defaults without saying more. What distinguishes the links connecting features of a default from other links in a predicative network? If nothing sets these apart, radical variability returns, and, at that

point my view begins to resemble concept nihilism. If category representations are too variable to single out and share, why call them concepts at all?

This is a serious concern, but I think it can be answered. Working memory limitations prevent us from activating representations that are overly complex. Therefore, only small bundles of features can be active at one time. Features that have been frequently co-instantiated are likely to form such bundles. This is what happens in simple associative models of learning. Default proxytypes are just the most common bundles. They are statistical freebies. They do not depend on any special unspecified aspect of predicative linking.

Predicative linking may face other problems, however. Ryder raises a few. He begins by asking me to provide an account of which features are transferred when a predicative link is formed. If we link a beagle image to a network containing images of other dogs, we don't transfer information about colour, but we do transfer information about some behaviours (a disposition for playing fetch). Here's a stab at a principle: if a feature in the dog network is inconsistent with the features in the beagle representation, do not transfer; if a feature is consistent but has been only rarely observed, transfer only if it has been observed in dogs that look like beagles; if the feature reflects the central tendency of the dog network, then (be disposed to) transfer. This needs refinement, but it already leads to testable predictions.

Ryder has another worry. He points out that representations tend to have links to many long-term memory networks. This poses a serious question concerning conceptual individuation. Does a beagle representation represent beagles, on a given occasion, or dogs? Weinberg fears that proxytype theory cannot explain how we ever think about dogs as such. Two things can be said in response. First, I follow Hume in saying that we represent general ideas by using particular ideas generally. If a beagle representation is being predicatively linked to a file with general information about dogs, and features specific to the appearance of beagles are ignored, then it is serving as a dog representation. Second, a single representation might represent both dogs and beagles on a given occasion. This is a somewhat radical possibility, not discussed in the book. Philosophers tend to assume that representations must have univocal content, but that may be wrong. Perhaps univocality of thought may be a specialty of creatures who have mastered a language. When one represents dogs using an auditory image of the word 'dog' ambiguity is somewhat more constrained. A labeled image may represent dogs without representing beagles.

There is a third worry in Ryder's commentary, which is even more serious. I say concepts can be individuated by the networks to which they are predicatively linked, but how, he asks, do we individuate the networks? To do that we need what Ruth Millikan calls a *sameness marker*, some way of determining which representations belong together. The most natural proposal is that representations belong to the same network if they have the same intentional content. Unfortunately, this would be circular, because, as Ryder points out, I think the intentional content of a representation is a function of the network to which it belongs. He also points out that the sameness marker cannot be sameness of vehicular properties of representations; on proxytype theory, representations of a category are too variable.

I think standard invariantist theories of concepts have no advantage here. On the standard account, when we encounter two category instances, we represent the fact that they are the same by labeling them with the same invariant representation. But what makes two tokens of the same label count as the same? As Millikan points out, it cannot be their intrinsic properties. Identical labels can represent different things. So we need a method, which can be functionally characterised, of treating two intrinsically identical labels as labels of the same category. That method can be equally applied if labels are different (compare matching two tokens of the word 'gnu' and matching the word 'gnu' with the word 'wildebeest'). Both the invariantist and the variantist need predicative links. Predicative links are introduced for different reasons on different occasions. Similarity of inner representations may be sufficient on some occasions, but not others. Millikan's view is akin to this. Representations are linked together by mental acts of identification, rather than by identity of their intrinsic features. If this is right, then an invariantist view of mental representations has no advantage over a variantist view.

Millikan's views suggest an account of memory network individuation. When we form our first representation of a category, we store it in memory. On future occasions, we identify things as belonging to the same category, and we do this for a variety of different reasons. On many of these occasions, new representations are stored along with the original. This collection of stored representations, created through multiple acts of identification, constitutes a memory network. When a concept is constructed in thought, its identity depends on the network from which it was constructed and the links it has to that network. This story of concept individuation is still thin on details, but I see no reason to think it won't work.

DESIDERATA

In the book, I say that all theories of concepts should be assessed against seven desiderata. All the reviewers raise questions about these, but I want to focus on a line of objection pushed by Weinberg. The core of Weinberg's worry concerns a useful distinction between what concepts are (concept ontology) and what concepts do (concept use). Some of the desiderata I offer may fall on the ontology side, while others pertain to use. Philosophers tend to assume that concepts are, ontologically speaking, representations that allow us to think about things. This picture corresponds most to the desiderata of intentional content (aboutness) and compositionality (thought formation). These desiderata are often taken by philosophers to be ontologically constitutive, while the categorisation desideratum is not. Weinberg says that, in treating all desiderata alike, I unfairly dismiss philosophical theories that fail to explain categorisation. This is a central theme in my argument against atomism (the view that lexical concepts do not decompose into meaningful parts). If we privilege intentional content and compositionality, my argument against atomism has little force. Concepts can satisfy those desiderata without serving as the mechanisms of categorisation. Categorisation can be explained by appeal to beliefs and other complex cognitive structures that contain concepts but are not identical to concepts.

Psychologists would balk at this objection. 'Concept' is a technical term, and they use it as a name for the mechanisms of categorisation. For them, the categorisation desideratum *is* ontological. It is a definition of what concepts are. So we have two competing ontologies of concepts. Philosophers say that they are representations that serve as building blocks for thoughts, and psychologists say they are mechanisms of categorisation. Perhaps we should just say there are two kinds of entities, and give them different names. Or as Yarlett and Ramscar suggest, perhaps we should just drop some of desiderata as unattainable philosophical ideals.

Things could work out this way, but it would be very nice if a single entity could satisfy philosophical and psychological demands. While atomism cannot be rejected for its failure to explain categorisation, it can be rejected if a more parsimonious theory can explain categorisation *along with* the things that atomism explains. My strategy in the book is to show that atomists posit entities (language of thought labels) that are superfluous. Suppose a physicist discovers that two forces thought to be independent can be explained by a single force. We would have reason to take such a theory seriously. Parsimony arguments have even more force in psychology. If a mechanism that evolved for one function can do another, there would have been no selection pressure on the evolution of another

mechanism. If mechanisms of categorisation can serve as mechanisms for thinking about things, then dedicated thinking mechanisms would be unlikely to evolve.

The categorisation desideratum and the intentional content desideratum have been sought by practitioners in two different fields, but they may converge on a single psychological entity. In the book, I argue that the convergence really shouldn't come as a surprise. On leading philosophical accounts, intentional content is conferred by correlations between concept tokenings and instances of the categories they represent in the world. Such correlations must be achieved by mental mechanisms that reliably detect the presence of category instances. Those, of course, are the mechanisms of categorisation. Psychologists' concepts serve philosophers' ends.

Weinberg has another worry that needs to be addressed. He argues that, in failing to acknowledge ontologies that emphasise the relationship between concepts and thinking, I overlook a crucial desideratum. Concepts, he says, must be thought-friendly. One aspect of this requirement is that concepts must readily lend themselves to thought formation and inference. If concepts were symbols in a mental language, this desideratum would be easy to satisfy. We could explain thinking on the model of logic. But there is no logic of highly variable, perceptually derived representations. It is unclear how such representations can be used to construct thoughts with logical form or how they can remain stable enough to participate in logical derivations.

Weinberg is quite right when he says that I need a more thorough explanation of thinking. My discussion of compositionality is intended to serve as something like a thought-friendliness desideratum, but I say far too little about logical inference. I do not think the problem is insuperable. In the book, I talk about logical operations over proxytypes. These operations can be defined by functional roles that conform to standard truth-tables. We can also achieve a substitute for logical form by simulating the truth conditions of logical constructions. Suppose you go to a restaurant and form the belief that if you order an appetiser, you won't have dessert. This can be represented as a series of meal simulations, none of which have both appetisers and dessert. Related ideas lie at the heart of Johnson-Laird's mental model theory. I think this approach has great promise.

There is also a more radical strategy for coping with logical inference. Above, I said that mastery of public language may facilitate disambiguation of concepts. Public language may also serve as a useful inference tool. Sentences have logical form that is more explicit than imagery. There is overwhelming evidence that we engage in a steady flow of silent speech.

Linguistic inference in silent speech may be a useful way to discover the entailments of our thoughts.

I cannot pretend that these brief remarks will satisfy the reviewers. Their objections are subtle and penetrating, and each warrants a more detailed response. In addition, I have neglected many important objections that were made in passing. See Weinberg on electrons, Ryder on primitive features, Yarlett and Ramscar on the assumption that senses have proprietary codes. *Furnishing the Mind* needs a sequel to adequately address all of these concerns.

Department of Philosophy
University of North Carolina
Chapel Hill, NC 27599
USA

