

Empiricism Regained?

By Dan Ryder

Jesse J. Prinz, *Furnishing the Mind*. Cambridge, Mass.: MIT Press, 2002. Pp x + 358.
U.S. \$40.00 HB.

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 pretty 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 will 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 vs. situational vs. 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 recognizing 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.¹ 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 categorization (construed very broadly to include judgements 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 categorization (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 recognizing or identifying someone. As Ruth Millikan (*On Clear and Confused Ideas*, Cambridge, 2000) has persuasively argued, 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 recognizing 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 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 categorization in these terms (just 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 representation, simple or complex, to explain this ability, it must possess some feature, a “sameness marker” (Millikan, *ibid.*, 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. 161-2.) 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 categorization. *Contra* Prinz, atomism does better.²

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 (e.g. the edge-detectors on p. 274, where he 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.³)

The network individuation problem ramifies beyond Prinz's unifying desideratum of explaining categorization. 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, a concept that denotes Monarch butterflies must be a reasonably good detector of Monarchs, i.e. it must causally covary with them.⁴

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 covaries only with (context dependent) contingent *signs* of the kind (p. 169). What really causally covaries 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 covariation properties*, on pain of trivializing 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 252-3), 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 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 categorization.

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 St.

Bloomington, IN 47405-7005

¹ Prinz swiftly dismisses theories that deny concepts are mental particulars, *inter alia* those espoused by Peacocke, Dennett, Wittgenstein, and Brandom – rather too swiftly, perhaps.

² Prinz faults atomism for failing to explain categorization because structureless concepts cannot explain discrimination. Well, yes, but atomism just explains discrimination by appeal to structures *external* to concepts, e.g. beliefs.

³ 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.

⁴ The concept will also covary 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 covaries with, a concept represents its “incipient cause,” the thing that caused its creation.