CHAPTER 9

Epilogue

> … the great regions of the mind correspond to the great regions of the brain.
>
> – Paul Broca, "Remarks on the Seat of the Faculty of Articulate Language, Following an Observation of Aphemia (Loss of Speech)"

> … when you are a Bear of Very Little Brain, and you Think of Things, you find sometimes that a Thing which seemed very Thingish inside you is quite different when it gets out into the open and has other people looking at it.
>
> – A. A. Milne, *The House at Pooh Corner*

9.1 Introduction

The aim of this book has been to provide an account of cognitive ontology, but rather than try to give some kind of comprehensive tour of the real kinds of cognition – if that were even possible – my approach has been to select a limited number of (what I take to be) both significant and representative taxonomic categories and use them to illustrate and support a theory of cognitive kinds, or real kinds in the cognitive domain. The general theoretical framework of real kinds was sketched out in Chapter 1 and it was applied to a variety of case studies in successive chapters of this book. Some of these cases are household categories that will be familiar to laypersons and cognitive scientists alike, such as *concepts*, *innateness*, and (*episodic*) *memory*, but others are not so well-known, for example, *language-thought processes* and *cognitive heuristics*. This is to be expected, since the aim is to choose some categories that are good candidates for kinds as well as some that are not, and the latter are less likely to be widely known or recognized. The fact that I had to come up with a label for "language-thought processes" is itself an indication that it is not usually treated as a cognitive kind in its own right. Still, analyzing this category is not just

an exercise in tilting at windmills, since the ways in which it fails to be a cognitive kind are instructive and illustrate certain general principles for identifying kinds in the cognitive sciences. The same goes for the category of *cognitive heuristics*. Moreover, in these cases and some others, even when a particular category is not a good candidate for being a cognitive kind, there are others in the vicinity that are (e.g. *myside heuristic*).

There is a danger in a book like this, containing successive chapters tackling separate case studies, that no single argumentative thread will emerge and that there may not be enough continuity among the cases to unify the whole. I have already attempted to indicate links and parallels in previous chapters, when they arose, but in this brief epilogue, I will try to draw further connections between chapters, and highlight certain common themes that bear emphasizing. In particular, in what follows, I want to draw attention to three main lessons that arise from these case studies. The first concerns the etiological–environmental individuation of many cognitive categories (Section 9.2), the second has to do with broader ontological issues in cognitive science (Section 9.3), and the third has to do with reductionism and the status of cognitive neuroscience (Section 9.4).

9.2 Etiological–Environmental Individuation

The categories discussed in this book pertain specifically to human cognition, not to psychology broadly understood. There are commonalities among categories in the cognitive domain that distinguish them from, say, categories in the domains of perception and emotion. These commonalities are engendered by the fact that cognition seems to delineate a relatively "closed system" (in the sense of Chapter 1). There are surely dense connections between cognition on the one hand and perception and emotion on the other, but the division between them is not an artificial one.[1] The categories discussed in previous chapters all pertain to the cognitive domain, though some of them are also implicated in processes involving perception and emotion, not to mention processes that lie outside the psychological domain altogether. For example, perceptual states serve as inputs to the capacity of *episodic memory*, and there is some evidence that the *myside heuristic* is influenced by affective processes. There is also evidence that *innate cognitive capacities* have a genetic basis. But these cognitive kinds

[1] See Beck (2018) and Phillips (2019) for two recent philosophical proposals for drawing the perception-cognition boundary on principled grounds, involving (respectively) stimulus-independence and stimulus control.

can be investigated using proprietary methods distinctive of cognitive science, and the causal processes in which they feature can be demarcated from others for the purposes of many inquiries.

One aspect of several of these cognitive kinds that I have tried to emphasize throughout is their externalist aspect, or as I have put it, their etiological–environmental individuation. This cumbersome expression is a more accurate way of describing the taxonomic practices that I have in mind for reasons that bear repeating here. It is misleading to describe the factors with reference to which some cognitive kinds are individuated as "external" primarily because the term "external" is insufficiently descriptive and does not indicate the precise features of reality that serve to ground these cognitive categories. In many of the cases discussed, what matters most is etiology: the causal history of the entities in question. Moreover, the causal history in question sometimes pertains to the developmental history of the individual and at other times the history of the species or phylogenetic lineage. The "externalist" label is not an apt epithet for these features of reality, but I will use it here because it has become so entrenched.

Externalism is commonly regarded as an established thesis in the philosophy of mind (though it is not without its detractors), but I will try in this section to distinguish the view defended here from other varieties of externalism in the philosophy of mind and cognitive science. As is widely known, there are two major philosophical proposals that emphasize the role of the "external world" in understanding the mind. The first is the externalist thesis associated with the work of Putnam, which is colloquially encapsulated in the expression, meanings "just ain't in the head" (1973, 704). The second is the "active externalism" or "extended mind thesis" associated with the work of Clark and Chalmers (1998), which holds that mental states and processes are characterized by "reliable coupling" with features of the environment. Starting with the second thesis first, the environmental–etiological thesis that I have advocated is not just about causal coupling but about the very identification or individuation of cognitive kinds. I have tried to show that many features of cognition are individuated with reference to their environment or etiology. This claim is more far-reaching than "active externalism" since it says that when it comes to something like a state of episodic memory (say), it is not even possible to identify it as such without invoking its causal history. If causal factors in the past or current environment of the thinker determine whether a certain cognitive state is a memory state in the first place, then it follows that memory states do not supervene on the thinker narrowly conceived. Hence, this thesis is stronger than one that claims merely that cognitive states of the thinker are

causally coupled with such entities as notebooks and smartphones (Clark & Chalmers 1998; Chalmers 2008). Indeed, it means that at least for some cognitive kinds, it is not clear how to frame the thesis that "internal" states are coupled with "external" entities, since it is misguided to talk about cognitive states being internal in the first place. As for the second familiar externalist thesis, it is explicitly an individuative thesis, but it is usually restricted to meanings, concepts, or semantic content. Concurring with this claim, I have argued in Chapter 2 that externalism holds for *concepts* (though I think they are individuated on the basis of both externalistic and internalistic factors), but I have also put forward reasons for thinking that it applies to other cognitive kinds, such as *innateness*, *domain specificity*, *episodic memory*, and *myside bias*. Moreover, the arguments I have made rely not on intuitive judgments concerning the ascription of conceptual contents, but taxonomic practices in cognitive science. For example, when it comes to concepts, I have tried to present empirical evidence showing that much of the empirical work in developmental psychology individuates the concepts of concept learners (partly) according to their etiology in the natural and social worlds (see Section 2.4). Though his emphasis has been on perception, Burge (1986; 2010) has also argued that externalism applies to kinds in scientific psychology. Based on this claim, he has made the most explicit case for the conclusion that the taxonomic kinds of psychology may not coincide with the kinds of physiology:

> Without a set of physical transactions, none of the intentional transactions would transpire. But it does not follow that the kinds invoked in explaining causal interactions among intentional states (or between physical states and intentional states – for example, in vision or in action) supervene on the underlying physiological transactions. The same physical transactions in a given person may in principle mediate, or underly [sic], transactions involving different intentional states – if the environmental features that enter into the individuation of the intentional states and that are critical in the explanatory generalizations that invoke those states vary in appropriate ways. (Burge 1986, 17)

My claim is that the etiological–environmental individuation of cognitive kinds is not only directly supported by taxonomic practices in cognitive science, but is also more far-reaching than many externalists have tended to assume. Moreover, some of the implications of externalism do not appear to have been widely acknowledged. In particular, apart from occasional stray remarks like Burge's just cited, it is not generally appreciated that the environmental–etiological individuation of cognitive categories may scuttle neat matchups with neural categories.

So far, I have emphasized etiology in the ontogenetic sense, but phylogenetic etiology is also relevant to cognitive science. For instance, I argued in Chapter 5 that what makes something a system or capacity of episodic memory is at least partly its function or what it evolved for. This means that in deciding whether nonhuman animals have episodic memory, we would be guided at least in part by the etiology of the systems or capacities.[2] More blatantly, when determining whether a cognitive capacity is domain-specific or not, I argued that this can only be done against a background of selection history. According to the account put forward in Chapter 4, the domain specificity of a cognitive capacity can only be identified with reference to the function that it evolved to perform. At least, that seems to be the most defensible understanding of the cognitive kind, as it is ordinarily deployed in cognitive science. These cognitive kinds are individuated primarily by their phylogenetic etiology, though that is not the only factor used to identify them. For example, domain-specific cognitive capacities are ones that do not generalize beyond their proper function, which means that they are individuated with reference to their synchronic causal powers as well as their etiology.

Etiological individuation is not the only type of relational individuation in cognitive science; in some cases, the synchronic environment of the thinker also plays a role in determining the identity of the cognitive kind that is manifested. For example, the central property in the causal cluster comprising the kind *innate cognitive capacity* is *triggerability*, which makes implicit reference to the informational content of the stimulus in the thinker's environment. Additionally, in the case of the *myside heuristic*, it was argued in Section 7.4, that determining whether an instance of the heuristic is indeed a *bias* requires doing so against the background of a particular cognitive task or problem. Hence, if researchers want to distinguish a myside bias from a myside heuristic, they can only do so with reference to the synchronic context of the thinker.

In the case of *concepts* and in some other cases, I have tried to argue that the kinds involved are individuated in part according to their etiology and in part according to their causal powers. They can be thought of as hybrid kinds. Generally speaking, if an entity has a certain causal history, that does not thereby confer on it any synchronic causal powers. For instance, there is nothing in common to all the human babies born on

[2] That need not mean that the only episodic memory capacities are homologues; they may also be analogues, where the functions are individuated at least partly etiologically. See Section 4.5 for further discussion.

January 1, 2000 (that they do not share with babies born on other days). Yet, in many cases in the biological domain, entities that share a history do share many synchronic properties because they have been produced by a causal process or mechanism whose function it is to generate entities with certain properties. A subset of such processes can be characterized as "copying processes" (Millikan 1999), although they need not copy entities perfectly – nor might that be their function. The outputs of such processes are what she calls copied kinds, including biological species and industrial artifacts. This phenomenon seems especially significant in the realm of cognition, where etiology often goes hand in hand with informational or representational content. When it comes to concepts in particular, a great deal of energy has been expended by philosophers on cases in which causal history pulls in one direction while causal powers push in another. These are cases in which the content of a thinker's concept has the "right" etiology but "wrong" causal powers, or vice versa (e.g. the famous Twin Earth cases). What ultimately determines concept identity in such cases? I indicated in Chapter 2 that there may not always be a determinate answer to such questions. One way to think about these cases is to regard them as inevitable outcomes of the convoluted causal structure of the world, which includes causal patterns that combine synchronic and diachronic elements. In discerning these patterns, there are sometimes various ways in which taxonomic lines can be drawn while preserving causal structure. That is why, in one sense, and for some purposes, the Earthian and Twin Earthian share a concept, while in another sense and for other purposes, they do not. But rather than bifurcating concepts to account for bizarre cases, it seems more parsimonious for cognitive scientists not to multiply concept categories beyond necessity, and deploy a hybrid category that is sensitive to both causal history and causal power. I have argued that that is what many cognitive scientists investigating concepts tend to do, albeit implicitly. This is notably the case in developmental psychology where concept possession, concept sharing, and conceptual change are of paramount importance.

This approach to cognitive taxonomy is in keeping with the characterization of the inaptly named "computational level," proposed by Marr (1982). As understood in this book, and as argued in previous chapters (see especially Section 1.5), I take the study of cognition to be conducted largely within Marr's computational level of analysis or explanation. By contrast with an algorithmic or implementational theory, a computational theory asks: "What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?" (Marr

1982, 25) Thus, when it comes to vision, Marr (1982, 23) characterizes a computational theory as follows: "In the theory of visual processes, the underlying task is to reliably derive properties of the world from images of it." The corresponding task for the theory of conceptual processes might be characterized as: to reliably generalize from particulars in the world so as to enable recognition, categorization, inference, and action. These accounts are broadly functional in character, where function is understood both synchronically and diachronically, and where functions allude to the worldly task or problem that the cognizer is solving or has evolved to solve.

9.3 Ontological Categories

In previous chapters, I have conjectured that the overarching or higher-order ontological categories characteristic of cognitive science are: *individuals* (or *objects*), *states*, *processes*, and *capacities*. At least, these are the broad ontological categories that include the kinds that have been explicitly discussed in this book. That is not to say that we might not need to posit other higher-order ontological categories, such as events or dispositions, in cognitive science. But it does seem as though these four categories capture almost all the kinds discussed in previous chapters. In this list of broader ontological categories, I have not included the category *mechanism*, which is conspicuous mainly by its absence in previous chapters. Many philosophers would expect mechanisms to play a leading role in cognitive ontology, but I would argue that they are more prevalent in the neural sciences than in the cognitive ones proper. Some philosophers might object that this unnecessarily narrows the scope of the *mechanism* category and its associated explanations, but what seems valuable about the category of *mechanism* is that it identifies a set of distinctive phenomena that lie within the causal sphere. When "mechanism" is understood in such a way as to encompass causality in general, we appear to lose this added descriptive and explanatory content.[3] The type of ontological configuration identified by many philosophers who have written about mechanisms and mechanistic explanation paradigmatically involves structures with proper spatial

[3] For an authoritative treatment of mechanisms, but one that in my view over-extends the concept, see Glennan (2017). For example, he states: "What is the relationship between mechanisms and causation? Put briefly, it is just that causes and effects must be connected by mechanisms" (Glennan 2017, 145). For a related criticism of some recent work on mechanism, see Ross (2020, 143), who writes: "A philosophical account that collapses these distinctions [e.g. between causal mechanism and causal pathway], and uses 'mechanism' as an umbrella term for all causal concepts, fails to capture such distinctions and their role in describing and explaining biological phenomena."

parts or components, arranged contiguously, and engaged in activities that involve physical contact. This type of configuration is more characteristic of neuroscience than it is of psychology, whose kinds and explanations tend to be functional rather than mechanistic. Being pitched mainly at the implementational level, many neural kinds are often structurally defined, have relatively well-defined physical boundaries, contain proper physical parts, and interact largely by means of physical contact.

The identification of the higher-order ontological categories that comprise the cognitive domain is one payoff of adopting the perspective of real kinds for the study of cognitive ontology. Another advantage of a real-kind approach is that it can help expose ambiguities in some of our taxonomic categories, for example, *episodic memory*, which can refer to a kind of capacity but also a kind of state. This is not exactly news, since memory researchers have pointed this out in various contexts, but in the case of episodic memory, reflection on this distinction brings out the fact that the capacity is primary, since *states* of episodic memory are such as a result of being produced by the *capacity* of episodic memory. This allows us to avoid tricky questions about the precise synchronic properties that ought to be possessed by states of episodic memory, for example, how much inaccuracy to tolerate in an episodic memory state before ruling it not to be a memory at all. On this approach, an episodic memory state is simply one that is produced by the capacity of episodic memory. Another advantage of a real-kind approach is that by distinguishing superordinate and subordinate categories and kinds, it becomes clearer that, for example, *episodic memory* might be a kind but not the superordinate category *memory*, or that *myside heuristic* might be a kind but not the superordinate category *heuristic*. Moreover, this implies that category labels like "episodic memory" and "myside heuristic" should not be parsed as identifying genus and species, or genus and differentia. The perspective of real or cognitive kinds also makes clearer the relationship between the cognitive kind *concept* and the cognitive kinds corresponding to specific concepts (e.g. *APPLE*, *ORANGE*), which I argued, is similar to the relationship between the kind *species* and specific species kinds (e.g. *Panthera tigris*, *Drosophila melanogaster*), rather than the relationship between superordinate and subordinate kinds. In fact, it seems closer to the determinable-determinate relationship (like the relationship of *color* to *red*, or *red* to *scarlet*). This was used to justify thinking of specific concepts as cognitive kinds in their own right.

In previous chapters, *episodic memory* has been classified as a kind of capacity, and so have *innateness* and *domain specificity*. But it might seem as though they are kinds of capacity in quite different senses. If a cognitive

capacity has been identified as one of episodic memory that seems to be different from classifying a kind of capacity as innate. Some might say that the former is a true kind of capacity, in the sense that all such capacities, in humans, other animals, and alien life forms, would belong to a single kind, whereas the latter is just a *property* of cognitive capacities. But the alleged contrast is overblown. Compare the following claim: In biological taxonomy, if an organism is classified as a *tiger*, that may seem fundamentally different from classifying it as a *predator* or *carnivore*. But that judgment does not seem warranted, if these are all real biological kinds. A real kind, I have argued, corresponds roughly to multiple (causally connected) properties. Thus, innate cognitive capacities typically share a host of causally connected properties, as argued in Chapter 3, just as episodic memory capacities do. Some kinds of cognitive capacity may be more inclusive than others, but that is no reason not to judge them to be kinds of cognitive capacity nevertheless.

9.4 Reductionism and Cognitive Neuroscience

The idea that cognitive phenomena will correspond in a simple and direct way to neural ones has been aptly labeled the "Simple Coordination Thesis" by Sterelny (2003). According to Sterelny (2003, 7), that thesis says that "meaning is a specific connection property of the wiring-and-connection facts," and he advises us to take seriously the possibility that "the relationship between the two sets of facts is much less clean" than the thesis supposes.[4] This is not to say that neural entities do not have representational properties, just that their representational properties need not translate directly to those of the full-blown cognitive or psychological entities that are associated with them.[5] Throughout this book, in examining taxonomic categories from the cognitive sciences, I have been guided by Marr's distinction between levels of analysis, and have maintained that the computational level is the proper domain of the cognitive. But

[4] At the risk of caricature, one can compare the Simple Coordination Thesis to the tendency of seventeenth- and eighteenth-century chemists, who subscribed to the atomic theory, to project from macro-properties of substances to their micro-constituents. For example, Whewell criticizes the French chemist Nicolas Lemery, who subscribed to the idea that acids taste sour because the micro-particles that constitute them have sharp edges, quoting him as writing: "I hope no one will dispute, seeing every one's experience does demonstrate it: he needs but taste an acid to be satisfied of it, for it pricks the tongue like anything keen and finely cut" (cited in Whewell 1840/1847, 382). The claim struck Whewell as both "gratuitous" and "useless."

[5] See Shea (2018) for a recent account of the representational properties of neural entities, which does not, it appears, presuppose something like the Coordination Thesis.

this might not square with a common conception of cognitive science, according to which it comprises such disciplines as neuroscience, or at least subdisciplines such as cognitive neuroscience. In fact, by thinking of the cognitive as a relatively "closed system," I might be accused of denying the very possibility of cognitive neuroscience. That would be an alarmist conclusion for two main reasons. The first is that even though I have argued that there are principled obstacles to reducing many cognitive kinds to neural kinds, that is not to deny that there are at least some reductions in the offing. The argument made in this book against blanket reductionism does not preclude the internalist individuation of *some* cognitive kinds and their identification with neural kinds. Second, and more importantly, just because there are not likely to be many reductions among psychological and neural categories, that is not to say that there are no important links between the respective sciences, indeed even among their categories. As many philosophers of science have noted, scientific disciplines and subdisciplines are related in intricate ways, notably by means of "interfield theories" that do not entail a direct or indirect reduction between them (Darden & Maull 1977). This results in a rather messier picture of the relationship between psychology and neuroscience than tends to be assumed in cognitive neuroscience and related subfields (cf. Stinson 2016). Indeed, it may have a salutary effect on cognitive neuroscience not to orient around the search for *neural correlates* of cognitive constructs. So far, that approach has not reaped dividends and it narrows the scope of expected relationships between the various research programs that investigate the mind–brain.

If neuroscience should not expect to find neural correlates for cognitive categories, how are we to proceed to build bridges between cognitive science and neuroscience? Specifically, what happens to cognitive neuroscience if investigators are not supposed to engage, for example, in reverse inference, which presumes that a particular cognitive capacity is being deployed on the grounds that its associated brain region is active? If there are no neat matchups between cognitive phenomena and neural ones, such inferences would appear to be ruled out and this methodological strategy would need to be revised. But, as I have tried to point out, such matchups are by no means always out of reach. It is just that many of the cognitive kinds investigated here cannot be expected to correlate with neural kinds, and they do not do so for principled reasons that seem to hold for a wide range of kinds in the cognitive domain. Furthermore, this outcome is to be expected given the precedent of other closely related pairs of sciences. Despite the intricate and intimate relationships between ecology and

genetics, to take a similar pair, there is no expectation that their considered categories will enter into a one-to-one correspondence. To be sure, there is a subdiscipline of ecological genetics, but it does not revolve around the search for genetic correlates of ecological constructs.

It may be objected here that the reason that the cognitive categories encountered in this book may not seem to correlate neatly with neural categories is that they are too mired in a prescientific and unsystematic way of thinking about cognition, including some of the relatively novel categories discussed in this book. Francken and Slors (2014; 2018) make a distinction between "commonsense cognitive concepts (CCCs)" and "scientific cognitive concepts (SCCs)," pointing out that there can be a lack of congruence between them. If many of the categories that cognitive science currently deploys are either identical with commonsense ones, or closely associated with them, we might do well to adopt a wait-and-see attitude before according them their own proprietary domain and declaring their relative autonomy from the implementational level.[6] Moreover, if the categories of cognition do not appear to be converging with those of neuroscience proper, that may be an indication that the former should be abandoned rather than retained.

I would provide two responses to this concern. The first is that even though many of our current cognitive categories may be revised, some of their distinctive features are likely to survive. The environmental–etiological individuation of cognitive kinds has been justified in large part by a certain approach to cognition pioneered by Marr. From this perspective, or so I have argued, the domain of the cognitive incorporates the developmental history of the organism and its phylogenetic lineage, as well as its current environment, in the investigation of cognizant behavior, and indeed in the very individuation of cognitive kinds. This aspect of the methodology and theoretical underpinnings of the study of cognition, is likely to survive the specific categories that cognitive science has devised. The second response speaks more directly to the indispensability of the cognitive domain. It is possible that many of our cognitive categories, particularly those that

[6] Indeed, it may be added that we have already begun to speak at least partly in neural terms and to replace some psychological terms with neural ones. However, many if not most such uses in ordinary parlance are clearly unwarranted. Francken and Slors (2018, 70) present interesting examples of misguided cases of "folk neuroscience" (e.g. "the surge in my endorphins was so swift and high that … I would lose all control"). In my view, there are at least a couple of things wrong with such locutions. First, those making such claims usually have no way of knowing that they are true (e.g. that endorphins actually surged). Second, they often mix and match neural and psychological terms without warrant, since the relationship between neurophysiological and mental phenomena is barely understood at this point.

originate in our folk theories (e.g. *concept*, *innateness*, and *memory*), may need to be discarded to keep up with the inexorable march of science. This may also apply to categories of more recent vintage, which may have been introduced with insufficient warrant (e.g. *domain specificity*, *myside heuristic*, and *Body Dysmorphic Disorder*). Ontological revisionism is a live option when it comes to cognition, as in other domains. Having said that, it cannot be denied that the relationship between our everyday taxonomic categories and those of a science of the mind are likely to be more intimate than with other sciences. In other domains, we can jettison taxonomic categories with abandon. But that is not the case when it comes to psychology or cognitive science. Much of our interest in developing a science of the mind lies in the ability to provide insights into the nature of our mental life. This means that if the categories developed by a mature cognitive science are entirely disjoint with our folk categories, this would not only weaken our interest in developing such a science but it would seem to preclude our using it to help us explain and make sense of our thoughts and actions as we ordinarily conceive of them. This is a common theme in the debate surrounding eliminativism about mental categories: that in eliminating our ordinary mental categories entirely, we would lose the ability to better understand ourselves (see e.g. Baker 2011). It is not inconceivable that we might move beyond the categories that we currently use to explain and predict people's actions and utterances, and that we would largely replace them with a new cognitive lexicon. But the prospects of doing so, at least in a thoroughgoing fashion, seem quite remote. For one thing, as I already indicated elsewhere (see Section 1.5; see also Section 2.4), wholesale conceptual revision is much less common in intellectual history and the history of science than is often supposed. For another, theoretical advances are often made by introducing new categories alongside current ones, or by dividing our current categories into sub categories while preserving the original categories. In other words, a complete overhaul of our cognitive categories is not in the offing, though some conceptual revision in light of empirical inquiry is certainly not to be ruled out.