CHAPTER 3

Innateness[1]

> Every human brain is born not as a blank tablet (a tabula rasa) waiting to be filled in by experience but as 'an exposed negative waiting to be slipped into developer fluid'.
>
> – E. O. Wilson (quoted by Tom Wolfe)

> Nature … is what we are put into this world to rise above.
>
> – Katharine Hepburn in *The African Queen*

3.1 Introduction

Innateness is somewhat unusual as a candidate for a cognitive kind since it may be more naturally considered a property than a kind. However, as I argued in Chapter 1, the distinction between properties and kinds is not a principled one and does not run very deep. Though kinds generally consist of clusters of properties, kinds can themselves sometimes be conceived of as properties. For example, the kind *Panthera tigris* is characterized by a number of etiological and causal properties, including the property of being a carnivore, but *carnivore* can also be considered a kind with individual members, such as members of the species *Panthera tigris*. The same would seem to go for *innateness*. It can be considered a property of particular cognitive capacities, but it can also be considered a kind, which includes particular cognitive capacities among its instances. In this discussion of innateness, I will focus primarily on cognitive capacities because I think that these are the most likely candidates for innate cognitive entities. In this context, cognitive capacities can include such things as the mindreading capacity in humans and birdsong in some species of birds. Therefore, in what follows I will focus on these kinds of cognitive capacities in attempting to justify the claim that innateness is a real cognitive kind.

[1] This chapter is an updated and modified version of Khalidi (2016a).

To say that innateness is a real kind is not to imply that any particular cognitive capacities or kinds of capacity are in fact innate. But it does mean that there is a meaningful question to be asked about whether a given cognitive capacity is innate. It may be that the extent of the innate human cognitive endowment is considerably less extensive than traditional nativist philosophers believed or many contemporary cognitive scientists posit. Nevertheless, I will argue that there is a non-vacuous kind, whose members can be distinguished from nonmembers. Another feature of the innateness category is that it appears to admit of gradations. It seems at least coherent to posit that some cognitive capacities are more innate than others. Unlike some other cognitive kinds, whether something belongs to the kind *innateness* might seem to be a matter of degree, as when we say that the capacity for birdsong in one species of bird is more innate than in another (though they both possess *some* degree of innateness). This may seem to rule it out as a kind-category, since kinds might be thought to be distinguished by an "unfathomable chasm" rather than a "mere ordinary ditch," to use Mill's expressions (1843/1882, 152). But I will try to show how these differences of degree can be understood on the model of innateness that I will propose, and will also try to justify the claim that this does not jeopardize the kindhood claim.

Though it originated as a folk or vernacular category, innateness has featured prominently in contemporary controversies in cognitive science.[2] There are ongoing debates concerning whether the linguistic faculty is innate to the human species, the extent to which numerical, spatial, and causal cognition are innate, and the degree of innateness of moral rules and religious concepts, among others. Yet, in addition to these first-order debates about the innateness of some of our basic cognitive capacities, there is a second-order debate about the viability of the innateness concept in the first place. Some scientists and philosophers have voiced reservations about the very category of innateness, questioning its suitability for rigorous scientific theorizing and doubting that it corresponds to a real cognitive kind. If they are right, then that would mean that the debates about the nature and extent of our innate cognitive endowment are pointless, since they revolve around a discredited concept. In response, I will attempt in this chapter to defend the concept of innateness against its critics and to establish that it is indeed a real kind in the cognitive domain.

[2] The OED defines innate, as applied to "qualities, principles, etc. (esp. mental)," simply as "opposed to acquired." The earliest usage cited is from Thomas Hoccleve, *The Regiment of Princes* (c. 1420), who speaks of "innat[e] sapience [i.e. intelligence]."

To appreciate the objections to the concept of innateness, it is necessary first to be acquainted with some recent efforts to define or characterize the concept of innateness. What is it for a cognitive capacity to be innate? Various attempts have been made to provide an explication of innateness that accords with contemporary scientific theorizing. Recent accounts of innateness include the following:

- *Canalization:* Innateness is canalization, where a phenotypic endstate is canalized to the degree to which the development of that endstate is insensitive to a range of environmental conditions under which the endstate emerges (Ariew 1999; 2007; Collins 2005).
- *Entrenchment:* Innate traits are ones that are generatively entrenched, in the sense that they appear early in development and have a number of later developing traits dependent on them (the degree of entrenchment is correlated with the number of traits depending on them) (Wimsatt 1999).
- *Psychological primitiveness:* Innate traits are explanatorily primitive in the domain of psychology, that is, the proper explanation of their acquisition lies outside of the domain of psychology (Cowie 1999; Samuels 2002).
- *Triggering:* An innate cognitive capacity is one that has a disposition or tendency to be triggered on the basis of an environmental input that is informationally impoverished by comparison to the resultant cognitive capacity (Stich 1975; Khalidi 2002; Khalidi 2007).
- *Closed process invariance:* An innate trait is one that develops across a range of normal environments and the proximal cause of its development is by a closed process or processes, where a closed process is one that tends to produce one or very few outcomes (and where closure is a matter of degree) (Mallon & Weinberg 2006; Weinberg & Mallon 2008).

Some of these proposals attempt to account for uses of innateness in a wide array of sciences, including microbiology, genetics, evolutionary biology, and so on, while others are intended to be more restricted to cognitive science. In what follows, I will be concerned with trying to rehabilitate a concept of innateness that applies primarily to the cognitive sciences. Rather than trying to characterize innate traits or phenotypic endstates in general, I will attempt to explain what it is for a cognitive capacity to be innate. As will become clear in due course, the characterization applies primarily to representational or information-bearing cognitive states. So it is quite possible that the innateness category has outlived its usefulness in many areas of biology, yet continues to be of value in cognitive science.

3.2 Critiques of the Innateness Category

All the attempts to elucidate innateness mentioned in the previous section have been deemed beside the point by other theorists on two principal grounds. The first criticism of the innateness category has to do with its alleged association with discredited essentialist and typological ideas in biology, while the second criticism claims that the category of innateness runs together a number of unrelated features or properties. Though these critiques are often combined, I think that they are distinct and will only focus here on the second criticism.[3]

The second criticism of the category of *innateness* charges that, rather than being a unified category, *innateness* is a multivalent category, combining a number of disparate criteria (Griffiths 2002; Bateson & Mameli 2007; Mameli 2008; Mameli & Bateson 2011; Shea 2012a). Closely related to this criticism is the contention that the different factors that influence people's judgments about innateness are not always correlated and that they can therefore lead to unwarranted inferences from one of those features to one or more of the others. The objection here is not that the features themselves are problematic but that the features, some of which are scientifically respectable taken individually, are not always correlated. The natural response to this objection consists in pointing out that innateness may be a polythetic category or a cluster category, which combines several of the features posited by the analyses mentioned and that these features are imperfectly correlated. By contrast with a monothetic category, the features are not singly necessary and jointly sufficient for the application of the category. Many other scientific categories, especially biological ones,

[3] I will not try to respond here to the first criticism, which has been made in a number of papers (Griffiths, Machery & Linquist 2009; see also Griffiths & Machery 2008, Griffiths & Stotz 2008, Linquist, Machery, Griffiths & Stotz 2011). But it is worth pointing out that even if the vernacular concept of innateness has some of the associations that these critics have identified, that may not render it unfit for scientific theorizing. Many scientific concepts originate as folk concepts before being refined and revised in order to make them suitable for scientific theorizing. In a recent empirical study of the innate concept among laypersons and scientists, Knobe and Samuels (2013) argue that members of both groups can engage in a process of "filtering" tainted concepts, dissociating them from unwanted prescientific associations. Hence, even when vernacular concepts have been implicated in prescientific or discredited scientific theories, scientists (and the folk) are capable of jettisoning their problematic features, especially when thinking explicitly and making considered judgments (as opposed to making snap decisions under time constraints).

are of this kind (e.g. biological species). Since scientific categories often appear in exception-prone empirical generalizations, there is usually a risk of making an unsuccessful inference from one of the features associated with a scientific category to another.

But the claim has also been made that the innateness category consists of a "clutter" rather than a cluster of criteria (Bateson & Mameli 2007; Mameli 2008; Mameli & Bateson 2011). Though these critics do not give a precise way of distinguishing a clutter from a cluster, they have likened the category of innateness to the pre-theoretical category JADE. As is now well known, jade comprises two distinct minerals, jadeite and nephrite, with entirely different chemical compositions. These substances share some of the same macro-properties (e.g. color, hardness) by sheer coincidence, but there is no superordinate kind *jade* that unites them. To escape the fate of jade, these critics state that a cluster theorist "has to give an account of the … properties that constitute the cluster, of the causal processes that connect such properties and cause them to tend to co-occur" (Mameli & Bateson 2011, 441). These researchers argue that some of the properties associated with innateness are scientifically sound but they are not always reliably correlated or causally linked. In what follows, I will try to respond to the challenge posed by the critics of innateness and will both outline the main properties associated with innateness and give a preliminary account of some of the causal connections among these properties. These properties are conceptually distinct and the causal connections among them may not be obvious, but I will maintain that there is considerable empirical evidence to suggest that they are reliably causally linked, enough to provide grounds for thinking that innateness constitutes a cluster category rather than a clutter.[4]

3.3 Innateness as a Cluster Category

A cluster category, with non-strict causal relations among the properties included in the cluster, is just what we would expect in a theoretical category in the biological and cognitive sciences. The question is, what are the associated properties in the case of *innateness*, and how are they related? We have already encountered some of them, and others can be readily gleaned from recent research in cognitive science. What follows is a tentative list

[4] A similar proposal has been made by Samuels (2007), but the features he associates with innateness are somewhat different from those I posit, and he continues to consider psychological primitiveness to be the primary feature associated with innateness, as I will go on to argue.

of properties or features associated with innate cognitive capacities, which may not be exhaustive:

- *Triggering* (or more properly, *triggerability*)*:* can be acquired in conditions of relative informational impoverishment
- *Lack of learning:* need not be acquired as a result of processes such as inference, conditioning, association, exploration, experimentation, repeated observation, and imitation
- *Early onset:* is acquired relatively early in ontogeny
- *Invariance:* is acquired across a broad range of environments
- *Canalization:* is buffered against environmental variation
- *Pan-cultural:* is present in all cultures, even though it may not be universal or monomorphic
- *Informational encapsulation:* is insulated from other cognitive content, functions independently of other cognitive systems
- *Cognitive impenetrability:* resists modification by other cognitive capacities
- *Critical period:* is acquired only or most effectively within a developmental window

In this section, I will try to show that, at least according to some prominent research programs in cognitive science, there are robust causal links among these properties. Cognitive scientists regularly make inferences and provide explanations that posit causal relations that situate these properties in a web or network, with some of these properties being causally prior to other properties in the network. Although I will not offer a comprehensive account of this causal network, I will try to provide enough examples to make a plausible case for the existence of such causal connections. Many of these examples will be drawn from a prominent nativist research program in developmental psychology, the "Core Cognition" research program championed by Carey, Spelke, and others. This is obviously not the only research program in cognitive science to use the category of innateness, but it is one of the most prominent ones to do so and at least some of its results concerning the development of human cognition have been widely accepted. It is therefore convenient as a relatively established scientific paradigm that has made widespread use of the category of innateness.

Triggering and Lack of Learning: If a cognitive capacity can be acquired on the basis of a trigger, that is to say that it can be acquired as a result of an impoverished input. While it may not make sense to say that an input is impoverished in an absolute sense, it is possible to rule that an input is impoverished relative to the resulting output, the output being the cognitive state or capacity that is acquired as a result of that input (Khalidi 2002; 2007). One of the most widely used argumentative strategies in cognitive science, the argument from the "poverty of the stimulus," relies on just such considerations to reach the conclusion that some particular cognitive state is innate (Laurence & Margolis 2001). But if something is disposed to be acquired as a result of an impoverished input, then it is plausible to say that it need not be acquired as a result of a learning process on the part of the agent, including trial-and-error, extensive observation, imitation, inference, or "any process that treats information derived from the world as evidence" (Carey 2009, 453). The reason is that treating the world as evidence involves an inferential process wherein information is processed and this generally involves something more than just minimal input from the environment. To be sure, the difference between learning something on the basis of information derived from the environment and having something triggered by an environmental input may be a matter of degree. Yet, learning and triggering can be thought to represent two ends of a continuum. To the extent that a cognitive capacity is disposed to be triggered, it can be acquired without learning. That is why some cognitive scientists consider that when a cognitive capacity has been acquired on the basis of an impoverished input, that is evidence that that capacity has not been learned (and hence that it has a substantive innate component). As Carey (2009, 196; emphasis added) puts it in discussing infants' capacity to represent agency: "In some cases the age of the infants, *along with considerations of limitations on the inputs they could possibly have experienced*, casts doubt on some plausible learning accounts." If one thinks of a trigger as an input that is impoverished relative to the resulting output, then it will lead to the acquisition of a cognitive capacity without the need for learning. Many research programs in the cognitive sciences treat innateness and learning as complementary notions, and though some also countenance degrees of innateness, it is safe to say that there is a negative correlation between degree of innateness and amount of learning.

Lack of Learning and Early Onset: If a cognitive capacity does not need to be learned then it can be expected to be acquired relatively early in development. If the acquisition of a cognitive capacity does not involve a process of deriving evidence from the world, then there is less of an obstacle to acquiring that capacity early in development. Developmental psychologists often argue from the early onset of certain cognitive capacities to the conclusion that they have not been learned, at least given other circumstantial evidence. For example, Carey (2009) holds that one of our core cognitive capacities includes a capacity to recognize others as agents and to represent others' mental states. Moreover, she thinks that this capacity of "core cognition" is not learned partly on the grounds of early onset and the lack of sufficient perceptual input. As Carey (2009, 196) puts it in a passage already quoted above: "In some cases the age of the infants … casts doubt on some plausible learning accounts." Even though the causal link from innateness to early onset may be weak, inference in the other direction is stronger, since it is unlikely that a cognitive capacity could be acquired so early in development by a process of learning. In other words, there is a causal link between lack of learning and early acquisition and there is also an inference to the best explanation from early acquisition to lack of learning. This causal link is taken for granted in so much recent psychological research in the past two decades that it explains the explosion of experimental work on infants younger than six months old, as well as the development of elaborate techniques for appraising the mental states of such preverbal and unreliable participants (most famously, the "violation of expectation" experimental paradigm).

Triggering and Invariance: Even though there are crucial exceptions, there is a natural link between triggering and invariance. If a capacity is capable of being triggered, manifesting itself on the basis of a relatively impoverished input, then it is likely to emerge in a variety of different circumstances and will therefore be relatively invariant across a range of environments (cf. Khalidi 2007, 109). This causal relationship can be illustrated by various types of studies, but it is perhaps most clearly revealed in research on animal cognition in which animals are reared in a variety of environments. For obvious ethical reasons, these types of experiments, which involve raising infant animals in a range of different environmental settings, are conducted primarily on nonhumans. The idea behind these experimental manipulations is to determine whether a cognitive capacity is relatively invariant across environments. If a capacity emerges in all or a wide range of different environmental contexts, this is regularly taken to show that it does not require specific experiences with the environment, and hence that it is merely triggered by the environment. There has been considerable research into spatial orientation in a variety of animal species, specifically regarding whether they use geometric or nongeometric information to orient themselves in unfamiliar environments, or to reorient themselves having been disoriented in familiar environments (see e.g. Hermer & Spelke 1996). In nonhuman animals, this aspect of spatial cognition is often investigated by raising animals in environments with a variety of different spatial configurations to ascertain whether this leads to important differences in their abilities to spatially orient themselves. Consider recent research into spatial orientation in chicks, which concludes that the cognitive capacity to reorient using geometric information is largely innate. Chiandetti and Vallortigara (2008, 144) observe: "No differences between chicks reared in circular-, rectangular, or c-shaped cages were apparent in the ability to reorient using purely geometric information (i.e. in the absence of any featural cues)." Partly on this basis, they conclude that "the results reported here for chicks … suggest that animals encode geometric information in the absence of (or with minimal) experience of surfaces of different lengths connected together at right angles" (Chiandetti & Vallortigara 2008, 144). That is, they infer that the cognitive capacity is triggered based on invariance across a range of training environments, since there is a causal link between a capacity's being capable of being triggered and its emerging across a wide range of environments. Again, the *inferential* link between invariance and triggering rests on a *causal* link in the opposite direction.

Invariance and Canalization: A cognitive capacity that is invariant is not always canalized, but one prominent way of achieving invariance in a cognitive capacity is by means of the ontogenetic device of canalization. A cognitive capacity that is canalized is buffered against environmental variation and develops according to a relatively fixed developmental pathway or a small set of such pathways. It is "programmed" to proceed along a finite number of different "channels" in such a way that precludes intermediate or hybrid developmental outcomes. Invariance in a biological setting can be achieved by means of a process of canalization, as in the acquisition of birdsong in many species of birds. In bird species in which the development of song is highly canalized, being buffered against environmental variation, the outcome (acquiring the species-specific adult song) will be invariant across a range of environments. Invariance can be achieved de facto in other bird species, where acquisition is not canalized, but it is so achieved only if a wide range of environments will contain the input needed in order to lead reliably to acquisition. It is worth observing that canalization does not lead to invariance in the sense of constancy of outcome, much less constancy of outcome across all possible environments. Rather, canalization produces a relatively *limited number* of outcomes across a very *wide range* of environments (i.e. relative rather than absolute invariance). In this case, as in some of the subsequent causal links to be explored below, invariance is not strictly causally linked to canalization, but given certain background conditions and plausible assumptions about implementation in a biological system, one prominent way of achieving invariance is by means of canalization. Moreover, the *direct* causal link operates in reverse: Canalization of a trait causes that trait to be invariant across a range of environments. Thus, if a capacity is canalized, it follows that it will be relatively invariant. However, one way that natural selection has devised to secure the invariance of a trait is to canalize it. The adaptiveness of invariance when it comes to certain traits has (so to speak) led natural selection to canalize some of those traits. This is why invariance can be said to be both an effect and a cause of canalization. The causal link here is a functional one in the sense that canalization has the "function" of bringing about invariance. As in the case of other causal-functional links, the effect (invariance) raises the probability of the reproduction of the cause (canalization), which in turn produces (another token of) the same effect, in a positive feedback loop.

Invariance and Panculturality: There is reason to expect, and considerable evidence to suggest, that capacities that are relatively invariant across environments will emerge in all human cultures. A pancultural capacity need not be monomorphic in the species, since it is quite compatible with its arising in all cultures that it be polymorphic or indeed that is a rare trait arising in a few select individuals. Still, invariance across a range of environments suggests that it would be compatible with the whole range of human *cultural* environments. Features of our core cognition are regularly claimed to be impervious to cultural differences in this way. Consider, for instance, recent work on concepts of OBJECT and SUBSTANCE, as manifested in human participants in the United States and Taiwan, native speakers of English and Mandarin, respectively. Since Mandarin is a classifier language while English is a count-mass noun language, some research has suggested that speakers of the former will tend to assume that a new word refers to a kind of substance while speakers of the former will tend to assume that it refers to a kind of object. But Li, Dunham, and Carey (2009) claim that though such effects can be observed in linguistic tasks, they do not occur in at least some nonlinguistic tasks. They conclude:

> The distinction between object kind and substance kind is a central piece of our core ontology, integral to our ability to make sense of and navigate a complex and shifting world. As such, it may not be surprising that, far from being highly malleable, it should prove itself quite resistant to linguistic or cultural influence, part of the shared conceptual endowment of our species. (Li, Dunham, & Carey 2009, 518)

Whether or not they are right to draw this conclusion, it is clear that the authors think that their nativist account of ontological categories links invariance to resistance to cultural influence. The fact that this cognitive capacity is regarded as invariant or not malleable is taken to imply that it is pancultural.[5]

Canalization and Cognitive Impenetrability: If a cognitive capacity is canalized then it is buffered against environmental variation, and this will often imply that it is impervious to input from other cognitive systems or that it is cognitively impenetrable. A canalized cognitive system needs to be protected from being modified, which means that its representational content should be resistant to being overwritten by other systems. One way of achieving this is by making the system cognitively impenetrable. Carey (2009, 68) considers one of the main properties of innate "core cognition" to be that it "is never overturned or lost, in contrast to later developing intuitive theories …" Canalization may not always lead to cognitive impenetrability, but unless a canalized system is cognitively impenetrable then it may not be sufficiently buffered against information from the environment that contradicts the information represented in that system. It is plausible that any genuinely canalized cognitive capacity would be shielded from being altered in this way.

Canalization and Informational Encapsulation: Similar considerations suggest that canalization among cognitive capacities may also be achieved by means of informational encapsulation. A cognitive capacity that is informationally encapsulated is one that does not depend on other informational systems to perform its function.[6] If a cognitive capacity is canalized, then it may emerge regularly after other systems and in roughly the same order, but if that capacity is itself one that appears early in development, then it ought not to depend on more developed cognitive systems to perform its function, on pain of being inoperative. Such appears to be the case with the system of spatial orientation in humans, which is dependent only on basic perceptual information to perform its function, as opposed to "higher" cognitive functions. Hermer and Spelke (1996) found that while human adults use both geometric and nongeometric information to reorient themselves in a novel room after being turned around several times to disorient them, children (aged eighteen to twenty-four

[5] To say that the distinction between objects and substances is innate is not necessarily to say that the concepts OBJECT and SUBSTANCE are themselves innate. But the innateness of the distinction may facilitate the acquisition of the concepts.

[6] The terms "informational encapsulation" and "cognitive impenetrability" are sometimes used interchangeably, but I am using them here in different senses, based on the way that they appear to be used in the empirical research that I am relying on. It could be said that the sense of informational encapsulation at play here is a rough counterpart, in the cognitive domain, of the notion of "entrenchment" used by Wimsatt (1999), which was mentioned in Section 3.1.

months) rely only on geometric information. Even though this geometric capacity for reorientation is informationally encapsulated, in the sense that it is largely self-contained and independent of other cognitive systems, by the time humans reach adulthood, they are also able to use other, nongeometric information to reorient themselves. This encapsulation of the capacity for spatial orientation in very young children, as well as its task specificity, are taken by Hermer and Spelke (1996) to imply that this capacity is an innate module. As Hermer and Spelke (1996, 227–228) put it: "Both the task specificity and the relative encapsulation of the reorientation process suggest, to a first approximation, that children's reorientation process depends on a 'geometric module …'" Though they have not tested directly whether this geometric module would emerge in a range of environmental settings, Spelke and Newport (1998, 315) argue that this too is plausible, on the grounds that the developmental environment for the subjects in these experiments is different from that of their evolutionary ancestors: "Because the laboratory animals and American children in these studies have not spent their lives in outdoor environments where hills and valleys uniquely specify object positions, but rather in rectangular environments where many symmetries make geometry-based reorientation prone to error, it is likely that this process has been shaped more by evolutionary history than by learning." Researchers infer from the informational encapsulation of this capacity to its modularity, as well as its canalization and lack of learning.

Canalization and Critical Period: Yet another way of bringing about canalization is by means of a critical period.[7] A cognitive capacity that is subject to a critical period is likely to be canalized at least in the following sense. An organism that receives the appropriate input or inputs within the critical period is then sent along a specific developmental pathway, whereas one that does not receive the appropriate inputs fails to be launched on that pathway. The critical period can be thought of as the entrance to a developmental pathway, without which the organism fails to proceed along that pathway or "channel." In some instances, there may even be two or more types of input, which when received during the critical period, determine different developmental pathways. This is one way of understanding

[7] A distinction is sometimes made between a *critical period* and a *sensitive period*, the difference being that the former entails a sharp cutoff in the ability to acquire a cognitive capacity while the latter involves a more gradual decline in that ability. However, as numerous researchers have pointed out, there are few if any sharp cutoffs of this sort in cognitive development, so all such periods are more properly thought as sensitive periods. But since it is the more common terminology, I will use "critical period," with the caveat that this does not imply a sharp developmental divide.

language acquisition in humans on the principles-and-parameters account (e.g. Baker 2001). Depending on the type of input received from language speakers in their environment during the first few years of life (the critical period), the parameters are set in a certain way and human infants then go on to acquire the syntax of the appropriate language. Canalization occurs because the input received during the critical period directs one along one pathway or another (or in the case in which no input is received, fails to proceed along a pathway at all). This phenomenon is evident in research on the human phonetic repertoire. In many instances, young infants can discriminate a wider range of phonetic contrasts than are made in their native language. If their native language makes a certain distinction, they receive input that observes the phonetic distinction within the critical period and it is consolidated; if it does not, then they receive no such input and their ability to make the distinction is lost. The consolidation of some phonetic discriminations in turn influences the acquisition of yet others, in what has been described as "cascading" critical periods, each constraining and directing the next (Werker & Tees 2005). These successive critical periods can act as channels along which phonetic development takes place. (More will be said about critical periods in Section 3.4.)

These connections among the properties associated with innateness suggest a causal network in which the instantiation of some of these properties leads reliably, though not inevitably, to the instantiation of others (see Figure 3.1). Moreover, in some of these cases, the links are causal-functional, in the sense that the property in question does not just lead in a linear manner to the production of the effect. Rather, it does so precisely because it was selected to do so and in doing so raises the probability that the cause will be reproduced. Hence in some of these cases (indicated in Figure 3.1), there is a causal feedback loop between cause and effect rather than a simple linear relationship.[8] Complex causal interactions among the properties associated with innate cognitive capacities clearly signal a difference with a category such as *jade*, which is given as a paradigmatic example of a "clutter" category by Mameli and Bateson (2011). Unlike the macro-properties of jadeite and nephrite, which are accidentally correlated, the properties of innate cognitive capacities are causally connected in various intricate ways.

There are two ways in which this account of innateness admits of degrees.[9] The first is the one alluded to in the discussion of triggering and learning. These complementary properties are themselves dimensional. To

[8] I am grateful to an anonymous referee for Khalidi (2016a) for urging me to clarify this point.
[9] For a more detailed discussion of degrees of innateness, see Khalidi (2007, 102–109).

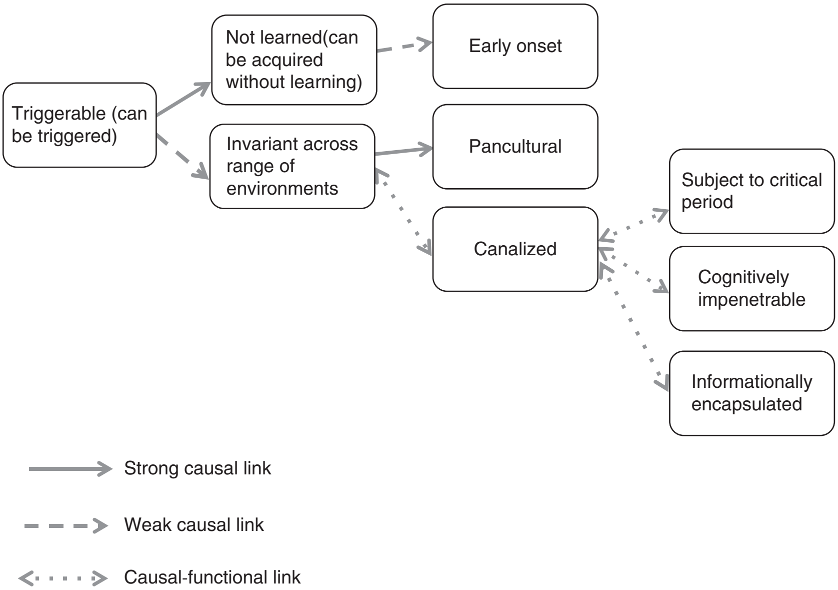

Figure 3.1. Causal network associated with the kind *innate cognitive capacity*

the extent that something is triggered, it is not learned, and this means that a capacity may be more or less innate depending (roughly) on the extent of triggering required to acquire it. In addition, the properties in this cluster are linked by non-strict causal relations. Hence, some instances of innate cognitive capacities may be characterized by more properties in the cluster than others, due to the operation of intervening causes. Those instances that are characterized by more of the properties in the cluster can be said to be more innate than others, even though they are both innate to some degree. This is simply a consequence of the fact that innateness, like many other cognitive kinds, is characterized by a cluster of properties that are linked by causal connections that are not strict. In fact, this is a characteristic of many paradigmatic kinds in the special sciences (and even, I would argue in the basic sciences), which is why such fuzzy kinds are ubiquitous in nature (Khalidi 2013, 65–69).

3.4 Is Innateness a Homeostatic Property Cluster?

Given the characterization of innateness in the previous section, it is natural to regard the category of an innate cognitive capacity as corresponding

https://doi.org/10.1017/9781009223645.004 Published online by Cambridge University Press

to a homeostatic property cluster (HPC), which conforms to the account of natural kinds developed by Boyd (1989), according to which natural kinds are clusters of properties kept in equilibrium by a certain causal mechanism. A similar proposal has been made by Samuels (2007), who explicitly identifies innateness as a HPC kind. As outlined by Samuels, HPC kinds satisfy three conditions: (i) They are associated with a number of features that tend to be co-instantiated, none of which is necessary for membership in the kind, (ii) a causal mechanism explains the co-instantiation of these features, and (iii) it is the causal mechanism, rather than the associated features, that constitutes the essence of the kind and defines membership in it (Samuels 2007, 23). There is a superficial similarity between my account of innateness as a real cognitive kind and Samuels' account, but I will argue in this section that my proposal differs from Samuels' in three important respects.

The first difference with Samuels' account concerns the particular features that I have proposed are associated with innateness. Though there is some overlap among the features we associate with the kind *innateness*, there is also some crucial divergence. Most significantly, Samuels considers that the central property of innateness, to which other properties are evidentially related, is psychological primitiveness, where a cognitive capacity is psychologically primitive if there is no psychological account of its acquisition. But far from being the central feature of innateness, I would argue that this is not a feature of innate cognitive capacities at all, since it has to do with the type of scientific explanation offered for the acquisition of an innate capacity rather anything about the capacity itself.[10] Moreover, several of the other features that Samuels associates with innateness, such as being *present at birth*, *adaptive*, and *monomorphic*, are ones that I do not consider to be associated with innateness with any regularity. Being present at birth is widely considered in the developmental literature not to be sufficient for innateness, since significant learning is now known to occur in the womb in humans and other organisms (see e.g. Partanen, Kujala, Tervaniemi, et al. 2013). It is clearly also not necessary, since many if not most innate characteristics, in cognition and elsewhere, are manifested well beyond birth, even in late adulthood. Being adaptive is also not strongly correlated with innateness; to think so is to commit a kind of adaptationist fallacy, since many innate cognitive and psychiatric disorders (though perhaps not all) are clearly maladaptive. Finally, monomorphism is no more

[10] I have put forward other criticisms of the primitiveness account elsewhere, see Khalidi (2007).

associated with innate features than polymorphism, since in humans as in many other organisms, there are a large number of significant dimorphic or polymorphic innate traits, from innate sexual traits to other phenotypic traits like eye color, hair color, blood type, lactose tolerance, and so on. It is likely that some innate cognitive features may be similarly polymorphic.

A second major difference with Samuels' proposal, is that he does not attempt to reconstruct the actual causal links that obtain between the features associated with the kind *innateness*, contenting himself with discussing "evidential relations" between these features. Although evidential relations are presumably grounded in causal connections, evidential relations can obtain between a cause and one of its effects, between an effect and one of its causes, as well as between two effects of a common cause. To say that features I_1 and I_2 are evidentially linked is to leave open the precise causal relationship between them. Instead, I have tried in the previous section to delineate these causal connections, at least in light of the current state of scientific research. While some of these causal links are straightforwardly linear, others appear to involve feedback loops, as mentioned in the previous section, and involve complex interactions between distinct properties. In order to justify the claim that innateness is a genuine cognitive kind, it is not enough to specify evidential or inferential relations among its associated features. One must show how these features relate ontologically to one another, as I have tried to do by means of causal links. This causal theory of innateness does not conform to the template outlined in Boyd's HPC theory of natural kinds, but it is consistent with the account of real kinds as "nodes in causal networks" that I explicated in Chapter 1 (see also Khalidi 2013).

The third difference with Samuels' account, and perhaps the most important, has to do with the fact that I deny the existence of a causal mechanism that keeps the cluster of properties associated with innateness in homeostasis, as specified by the HPC account of natural kinds. Whereas Samuels regards the "mechanism" as the causal essence of the kind, I think it is problematic to posit a causal mechanism that defines the kind or constitutes its underlying essence, at least if the term "mechanism" is used with anything like its standard meaning. Sometimes, the term "mechanism" is used so loosely that it is roughly synonymous with "cause," but there is a more precise usage that has become very widely accepted, according to which it refers to "entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions" (Machamer, Darden, & Craver 2000, 3). On Boyd's account of natural kinds, there is typically a causal mechanism that keeps the properties associated with any given kind in a state of equilibrium or homeostasis, understood roughly along these lines. In the case of innateness, a natural suggestion would be that the mechanism that keeps the properties in the cluster in a state of homeostasis is *genetic*, and that the entities and activities that produce the innate cognitive capacities are genetic in nature. However, there are two related problems with this suggestion. The first is that the genetic cause is likely to be so far upstream from the cognitive capacity that it does not seem possible, in general, to think of it as the single mechanism that keeps the properties in homeostasis. That is simply because there is a looseness of fit between the genetic factors that may contribute causally to the emergence of an innate cognitive capacity and that capacity itself, since there are numerous intermediate processes that may contribute to the outcome. As Spelke and Newport (1998, 291) once observed: "the central accomplishments of recent research in developmental neurobiology are to reveal a host of epigenetic processes through which neural structures develop in accord with a species-typical, intrinsic plan, without either shaping by the environment external to the organism or detailed genetic instructions." The second problem with positing a genetic cause common to all innate cognitive capacities is that it is unlikely that all innate capacities will be underwritten by the same type of genetic substratum. The point is not merely that there is likely to be a different genetic substratum that brings about the innateness of cognitive capacity C_1 and cognitive capacity C_2, but that the genetic substrata in question may not themselves share any important properties in common (beyond being somehow genetically instantiated). This is not to rule out that we may, in the case of some particular cognitive capacity, find that a certain genetic mechanism codes for that capacity. In the most idealized scenario, there may be a specific nucleotide sequence that codes for a protein, which then encourages the formation of synaptic connections between groups of neurons, which in turn are the neural substrate for a putatively innate cognitive capacity, such as the capacity to represent object permanence or analog numerical magnitude. But there may well be other innate capacities for which the causal story is far more complicated, involving multiple regions in the genome, regulatory mechanisms, epigenetic processes, interactions with the environment, and so on. An analogy with innate diseases may help: Consider two diseases that are largely innate, Huntington's chorea and cystic fibrosis, but for quite different reasons and as a result of very diverse causal pathways. Huntington's results from an abnormally long trinucleotide repeat near the tip of chromosome 4, while cystic fibrosis turns out to involve a more heterogeneous collection of mutations (Kitcher 1997, 60). Hence, what it is for a disease to be innate (or partly innate) can be the result of different types of genetic (and epigenetic) processes. The situation is likely to be similar in the case of innate cognitive capacities: What it is for each of them to be innate at the genetic level may differ considerably, so there is little prospect of locating a single type of genetically based mechanism that is common to the development of all innate cognitive capacities.[11] In the absence of a single type of genetic mechanism underwriting all innate cognitive capacities, the regularly co-occurring properties associated with an innate cognitive capacity may be understood not as having a single underlying cause, but rather as being themselves causally interconnected in various ways. That is (as Boyd himself acknowledges at times), there need not be a single mechanism to keep them in homeostasis.[12]

It may be objected here that the mechanism that corresponds to the innateness of a cognitive capacity need not be genetic but may instead be neural. That is, even though the genetic bases of innateness may be diverse, there may be a common neural mechanism or process that corresponds to the innateness of a cognitive capacity. Perhaps there is a certain type of neural profile that all innate cognitive capacities share, such as a certain mode of neural connectivity that renders them capable of achieving a mature state with relatively minimal external stimuli. It is certainly possible that there are neural commonalities corresponding to the cognitive kind *innateness*. But given that I have argued that the central property in the innateness cluster is the disposition to be triggered, it is at least possible that innate cognitive capacities are susceptible to triggering via different routes. The disposition to be triggered by external stimuli is a prime example of a functional property that can be realized differently in different physical structures. To use a simple analogy, the triggering mechanism of a gun is different than the tripwire mechanism that triggers a landmine. In both cases, the result is an explosive process that can be initiated by a relatively weak external stimulus, but this disposition is realized by very

[11] Shea (2013) has developed a notion of *inherited representation* that attempts to escape this problem of genetic and epigenetic heterogeneity. Without trying to rehearse this notion in any detail, it may serve to encompass a variety of different genetic and epigenetic mechanisms for encoding cognitive capacities so that they can be passed from one generation to the next. Thus, I would not rule out the possibility of a viable characterization of the mechanism that holds the features listed in homeostasis.

[12] "Either the presence of some of the properties in [a family of properties] *F* tends (under appropriate conditions) to favor the presence of the others, or there are underlying mechanisms or processes which tend to maintain the presence of the properties in *F*, or both" (Boyd 1989, 16).

different structural means. Even in the absence of direct evidence that a disposition to be triggered is multiply realized in the cognitive domain, it seems clear that it is multiply *realizable*. There may be different neural mechanisms leading to the development of a cognitive capacity with minimal triggering or in the absence of learning. Hence, it is at least possible for there to be a one-to-many relation between the cognitive kind *innateness* and its neural correlates. Moreover, it is also possible for the relationship to be many-to-one. Given the relational nature of some of the properties associated with innateness, the very same neural properties that underpin them may correspond to different cognitive kinds in different contexts. This appears to be true of the disposition to be triggered, which is characterized in terms of the relative informational contribution of the environmental stimulus and the resulting cognitive capacity. If triggering is assessed based on the informational content of the stimulus in relation to the resulting cognitive capacity, it is not clear how this could be captured purely in neural terms (even if we had the means to discern the informational content embodied in neural mechanisms). In addition, I have made a similar case elsewhere (Khalidi 2020) for the cognitive kind *critical period*, which is characterized in terms of its place in the lifespan of the organism. The same neural mechanism that correlates with a critical period may be correlated with a different cognitive kind if that mechanism were to operate toward the end of an organism's lifespan. Thus, at least some of the properties in the innateness cluster in the cognitive domain are individuated contextually with reference to environmental or developmental contexts. If in some cases the corresponding neural mechanisms are not so individuated, this would yield a many-to-one mapping between the cognitive kind and its neural correlates. The result is a many-to-many mapping between cognitive and neural kinds.

I have now expanded on three ways in which this account of the cognitive kind *innateness* does not conform to the standard characterization of HPC kinds, as clusters of properties kept in equilibrium by a causal mechanism. Instead, according to the position that I have outlined, *innate cognitive capacity* is a natural kind associated with a cluster of properties related to one another by intricate causal connections, though not held together by a single causal mechanism. This account also allows us to respond to a recent critique of the innateness category that has been put forward by Shea (2012b). One of his principal reasons for rejecting the proposal is that the clustering of properties associated with innateness is especially unreliable when it comes to human beings, mainly on the grounds that inherited representations are not necessarily genetic but can also be epigenetic, cultural and so on. The idea seems to be that some of the properties associated with innate cognitive capacities may be the result not of inherited *genetic* representations, but of (say) inherited *cultural* representations, and this may lead us to conclude falsely that a cognitive capacity is innate when it is not. To use an illustration of my own, suppose that the cognitive capacity to read and write is a human cultural invention that has been transmitted as a result of imitation. If we observe that reading and writing are pancultural and highly canalized, we might be tempted to infer that they are innate. But we would be rash to conclude this, since panculturality and canalization may in fact be the result of an inherited cultural representation and therefore not innate. If that is the worry about innateness, then it seems misplaced. Even if there are other causal properties that tend to generate some of the same properties in the innateness cluster, as I have identified them, that should not lead us to conclude that innateness is not a genuine cognitive kind, though it may make it harder to distinguish innate from non-innate cognitive capacities. There is a rough analogy here with the case of jade, considered earlier, but with the following crucial difference. The fact that the macro-properties of jadeite and nephrite largely coincide for accidental reasons tells against counting *jade* as a natural kind, but it surely does not undermine the case for considering either *jadeite* or *nephrite* separately as natural kinds. Just because many of the properties in the jadeite cluster coincide with those in the nephrite cluster, that should not deter us from thinking that either of them is a natural kind in its own right. Similarly, *innate cognitive capacity* may well be a natural kind even though some of the properties that are causally associated with it are also associated with cognitive capacities that are culturally transmitted and not innate.

In this section, I have tried to show that there are three important differences between this account and Samuels', differences that would rule out considering innateness a natural kind along the lines of Boyd's homeostatic property cluster kinds. Instead, innateness can be considered a natural kind along the lines of the simple causal theory of natural (or real) kinds defended in Chapter 1.

3.5 Objections and Replies

One objection to the argument that I have presented pertains to the very identity of the category that I am purporting to defend. The revamped category of innateness that I have tried to defend might be said to bear little resemblance to the original folk category of innateness, suggesting that we should scrap it and start anew rather than attempt to brush up a tired vernacular category. In this vein, Mameli (2008) compares the concept *innateness* to the concept *mass*, which the special theory of relativity allegedly eliminated and replaced with two concepts, *rest mass* and *relativistic mass*, neither of which coincides with the concept *mass* as found in classical mechanics.[13]

A response to this objection must address the question of conceptual change and continuity, as well as the relationship between scientific concepts and lay concepts (albeit briefly). The descriptive question as to when a concept has been retained and when it has been replaced, and the related prescriptive question as to when we ought to retain a concept and when to replace it, are controversial to say the least. But it seems safe to say that when there is significant continuity between a lay concept and a scientific concept, retention is usually the outcome. Moreover, there are two prominent prescriptive considerations that would tend to favor retention in many cases. First, an entrenched lay concept is sometimes difficult to abandon entirely and attempts to expunge it are often counter-productive and likely to be met with resistance. Second, as long as there is some continuity, it is more conducive to the comprehension and communication of scientific results to express them mainly in terms of existing concepts (wherever possible) rather than introduce altogether new concepts. These points can be illustrated using the concept HEAT as it was used by both scientists and laypersons in the western world until the mid-eighteenth century, that is until Joseph Black's proposal of the theory of latent heat. Until this time, scientists had not clearly differentiated the concepts of KINETIC ENERGY and TEMPERATURE (Carey 2009, 372; cf. Wiser & Carey 1983; Wiser 1988). Whereas kinetic energy is an extensive and additive quantity, temperature is intensive. If two cups of water at the same temperature are added together, the quantity of kinetic energy in the mixture is increased but the temperature of the water remains the same. Once this differentiation was made, all subsequent scientific theorizing on the subject proceeded to distinguish the two concepts. How should we describe this episode in intellectual history or the history of science: Is

[13] Mameli's claim concerning the concept MASS is controversial and seems to endorse a view according to which the concept did not survive the theory change from classical to relativistic physics, a view that has been widely disputed. Many scientists and philosophers of science have argued instead that REST MASS should be identified with the classical concept MASS and that the latter concept has not been eliminated at all (see e.g. Earman (1977), Earman & Friedman (1973)). Note also that it may be misguided to insist to an engineer that one should not talk about MASS, but must always distinguish REST MASS from RELATIVISTIC MASS.

it one of elimination or modification? We could say that the folk concept HEAT was abandoned in rigorous scientific thought and replaced by two distinct concepts, KINETIC ENERGY and TEMPERATURE. Alternatively, we could (and often do) say that the scientific concept KINETIC ENERGY is roughly equivalent to the vernacular concept HEAT, and that the latter concept has been modified to differentiate it from the concept TEMPERATURE. In the former case, we have eliminated a folk concept in favor of two scientific concepts, and in the latter case, we have modified a folk concept and rendered it roughly equivalent to a scientific concept. If our linguistic practices are anything to go by, in this case, it would seem that elimination has lost out to modification. The term "heat" in English, like related terms in other languages, survives today and is commonly regarded as a loose synonym for the scientific term "kinetic energy" in certain contexts. If that is the correct way to describe the situation, we can still ask a prescriptive question as to whether this was the right course of action, or whether it might not have been preferable to abandon the folk concept in favor of the two scientific concepts. Though it might still give rise to misconceptions among the folk and lead to occasional mistaken inferences, there are advantages to modifying an existing concept rather than eliminating it altogether. Again, there are two considerations that favor retention. First, since the concept HEAT was highly entrenched, it might not have been feasible to eliminate it altogether. Second, as long as there is some continuity, the sciences are often better served when they can relate their findings to our prescientific theories rather than when they introduce new jargon or specialized language that is not readily accessible to the lay public. We do not normally say that there is no such thing as heat, but teach schoolchildren that heat is different from temperature. Similarly, in the case of innateness, it might be more productive to frame our scientific findings in terms of innateness rather than try to purge it from our vocabulary. Moreover, there is sufficient continuity that it seems feasible here to express our scientific findings in terms of our existing concept. Rather than saying that there is no such thing as innateness, it may be more productive to point out, say, that while it is true that what is innate is not learned, it is not the case that what is innate is always present at birth.[14] This stance

[14] Recent detailed investigations of the interactions between lay concepts and scientific concepts indicate that the relationship is more complex than some philosophers have hitherto believed, and may not always involve deference by laypersons to the scientists. For instance, Radick (2012) relates that the late nineteenth- and early twentieth-century geneticist William Bateson resisted the concept HEREDITY but eventually succumbed to widespread usage, indicating that scientists sometimes defer to the lay public.

regarding concept individuation coheres with the account of concepts that I defended in Chapter 2. On that account, concepts are individuated on both externalist and internalist grounds. In this case, the scientific concept and folk concept apply to many of the same capacities (e.g. to reasoning, but not to reading), and there is considerable continuity in the functional-causal profile of the concept (e.g. the contrast with learned capacities, the association with early onset). Hence, there is sufficient reason to consider that the folk and scientific concepts coincide. (Recall also that this does not mean that there is exact coincidence between all the associated beliefs.)

Another, not unrelated, objection suggests itself here. If we attach the label of *innateness* or *innate cognitive capacity* to the entire causal network that I sketched out in Section 3.3 (and represented in Figure 3.1), it may be said that this cluster of causal properties corresponds more closely to the presumptive natural kind *cognitive module*, roughly in the sense first articulated in Fodor (1983), rather than the natural kind *innate cognitive capacity*. In other words, the kind I have described might seem to have a number of distinguishing features that go well beyond the bare kind *innateness*, even *innate cognitive capacity*.

There is likely to be some overlap between the kind *innate cognitive capacity* and the kind *cognitive module* (assuming the latter is also a cognitive kind), and the latter may indeed be a subordinate kind of the former. Though there may be some cognitive modules that are not innate, many of the cognitive modules now posited in cognitive science are in fact thought to be innate adaptations (though they may be far fewer in number than the hundreds posited by some evolutionary psychologists). Similarly, there may be some innate cognitive capacities that are not full-blown modules, but few philosophers and cognitive scientists nowadays think that there will be isolated innate concepts. If innateness pertains primarily to innate cognitive capacities, then I have argued that they will tend to have a cluster of features in common (e.g. canalized, not learned, informationally encapsulated, and so on), given what we know about human cognition. When Fodor posited modules, he also associated them with other properties that I have not associated with innate cognitive capacities (e.g. fast, automatic, mandatory), though perhaps a weak causal link with some of these properties may also be discovered. Be that as it may, there will be exceptions to most of these properties in any given instance and the clustering of these properties will not be perfect. In explicating the kind *innate cognitive capacity*, I have drawn the boundary around a range of properties that are often, but not always, co-instantiated, and one could draw it more strictly around a smaller subset of properties that are more strongly associated with one another. But then one would miss some causal processes and neglect a number of (exception-ridden) empirical generalizations.

A third objection would question the comprehensiveness of the account that I have given, asking whether there may not be other features associated with innateness that I have failed to consider. In response, it bears repeating that there are various alleged features of innateness that I have already argued do not pertain properly to it, for example, several of those mentioned by Samuels, such as being present at birth, adaptive, and monomorphic. But that obviously does not rule out the existence of other properties that ought to have been included in my account but were not. Mameli and Bateson (2006) list twenty-seven features commonly associated with innateness (in order to argue that none of them ought to be identified with innateness), but most of these features are either ones that I have included or ruled out. The latter category includes various properties that only pertain to the biological rather than the cognitive domain, or those involving genetic determination (e.g. being genetically encoded or genetically influenced), since I have argued that there need be no genetic mechanism that is responsible for the properties associated with innateness. There may yet be other features that I have not considered, which may prove to have causal links to the network I have elaborated. It is difficult to say with finality that no other properties are involved, but if there are other properties then I submit that that would tend to enrich the causal network and strengthen this account of innateness.

3.6 Conclusion

The category of innateness has been criticized as being incoherent and inappropriate for use in a mature cognitive science. In this chapter, I have tried to argue that at least some of the properties associated with the category of innateness are causally linked in a manner that is generally characteristic of real kinds. In Chapter 1, I argued that kinds in the cognitive sciences are validated by the role that they play in causal networks, like kinds in other sciences, both basic and special. When the instantiation of a property or, more commonly, the co-instantiation of a cluster of properties leads causally to the instantiation of a multitude of other properties in recurring causal processes, we identify such a property or set of properties with a natural or real kind. At the current juncture in cognitive science, innate cognitive capacities seem to fit this template well. I also argued in this chapter that there is not likely to be a single type of genetic or neural mechanism or process that underlies all innate cognitive capacities, which suggests that the cognitive kind *innateness* is multiply realizable relative to neural kinds. Moreover, given the nature of the properties associated with the kind *innateness*, some of which are relational or etiological in nature, there is also likely to be a many-to-one relationship between this cognitive kind and its neural correlates. This reinforces one of the central claims in this book, namely that in at least some cases, the relationship between cognitive kinds and neural kinds is likely to be many-to-many.