CHAPTER 4

Domain Specificity

> It is also very worthy of remark, that, though there are many animals which manifest more industry than we in certain of their actions, the same animals are yet observed to show none at all in many others …
>
> – René Descartes, *Discourse on Method*

> Evolution behaves like a tinkerer who, during eons upon eons, would slowly modify his work, unceasingly retouching it, cutting here, lengthening there, seizing the opportunities to adapt it progressively to its new use.
>
> – François Jacob, "Evolution and Tinkering"

4.1 Introduction

There are many contexts in cognitive science in which it is useful to distinguish domain-specific cognitive capacities from domain-general ones, and the distinction seems to do some explanatory work. For instance, according to many evolutionary psychologists, human cognition consists largely of domain-specific cognitive capacities, and this feature of our cognitive makeup provides evidentiary support for the pervasive influence of evolutionary processes on the formation of the human mind (e.g. Carey & Spelke 1994; Cosmides & Tooby 1994). By contrast, according to other researchers, domain-general cognitive abilities are the norm in the human mind and are what distinguish human cognition from that of most other animals (Samuels 1998; Fodor 2000). Yet, the distinction between the two kinds of capacities, domain-specific and domain-general, is not easily drawn. Moreover, some of the examples that are put forward to illustrate the distinction seem to be either spurious or misleading. Like innateness, domain specificity is commonly thought to be a feature of cognitive capacities, processes, systems, and related cognitive kinds. It tends to be applied to capacities that are restricted or constrained in their range of operation in some way. The challenge is to spell out the nature of this restriction and to demonstrate that all such restricted cognitive capacities have something substantive in common that would warrant treating them as members of a kind.

One of the main challenges in explicating the cognitive construct of domain specificity has to do with the difficulty of defining domains in the context of cognition. Researchers in this area regularly distinguish domains like: physical objects, language, number, morality, and theory of mind. But the boundaries of these domains are very permeable and it is not immediately apparent what it would mean for a cognitive capacity to be restricted to one of these domains, or indeed, what it would mean for it to be able to range across (or beyond) one of them. Thus, one of the tasks of this chapter will be to address this difficulty and determine whether it can be satisfactorily resolved. Another challenge is that the term "domain-specific" appears in a number of guises in cognitive science, and some uses seem more tenuous than others, or more aptly captured by other expressions. In what follows, I will focus on what I take to be the most influential usage, though I leave open the possibility that there are other uses that are also worth preserving (preferably using other labels).

In this chapter, I will begin in Section 4.2 by relating domain specificity to some other cognitive constructs, including the constructs of *modularity* and *innateness* (the latter of which was examined in Chapter 3). In Section 4.3, I will propose an account of the phenomenon of domain specificity based on a paradigmatic example as well as on existing theoretical proposals in the cognitive science literature. I will test this account of domain specificity on additional examples in Section 4.4, bringing out some of the distinctive features of this understanding of domain specificity, and attempting to determine whether the category of *domain specificity* corresponds to a cognitive kind. In Section 4.5, I will respond to a theoretical challenge to characterizing domain specificity, which has been termed the "grain problem," before concluding in Section 4.6 that domain specificity can be considered a cognitive kind when suitably described.

4.2 Domain Specificity and Its Confounds

Domain specificity is a feature of cognitive capacities that is often associated with several other such features, notably: modularity, innateness, and brain localization. In this section, I will examine the connections that may or may not exist between domain specificity and these other features, in order to gain a preliminary understanding of the cognitive category of *domain specificity*. By virtue of the way in which *modularity* was initially defined by Fodor (1983), there is a strong link between modularity and domain specificity. Indeed, it follows from Fodor's account that domain specificity is one of the defining features of modularity, and therefore that all modular cognitive capacities are domain-specific.[1] Of course, a case might be made for rejecting this definition on the grounds that it is unwarranted by the empirical facts or otherwise detrimental to research in cognitive science, but I will not try to make that case, nor do I think that the case can easily be made. Fodor's definition appears to have been widely accepted by cognitive scientists, including those who reject modularity, and I will not try to oppose it here. Hence, I take it as uncontroversial that domain specificity is one of the characteristic features of modularity, though the two concepts are not identical and ought not to be conflated.[2]

Things are more complicated when it comes to *innateness*. Although there is also a widespread assumption that there is a link between innateness and domain specificity, there is no prima facie reason for inferring such a link. It is not obvious that all innate cognitive capacities are domain-specific, nor that all domain-specific cognitive capacities are innate. To illustrate, human beings may have an innate cognitive capacity for associative learning that may be entirely domain-general. Conversely, there may be certain cognitive abilities that appear to be domain-specific that are not innate but

[1] In addition to being domain-specific, according to Fodor (1983), modular cognitive capacities are supposed to: (2) process items automatically and in a mandatory manner, (3) be inaccessible to consciousness, (4) be fast, (5) be cognitively impenetrable (e.g. resistant to being unlearned), (6) process "shallow" or highly salient features, (7) have fixed neural architecture, (8) have specific breakdown patterns (as in aphasia, agnosia), and (9) have fixed ontogeny (standard pace and sequence of development). Many subsequent discussions take domain specificity to be one of the most central features of modularity. For example, Sperber (1994, 40; emphasis added) writes: "The rough idea of modularity is also clear: A cognitive module is a genetically specified computational device in the mind/brain (henceforth: the mind) that works pretty much on its own on *inputs pertaining to some specific cognitive domain* and provided by other parts of the nervous systems (e.g., sensory receptors or other modules)."

[2] For the sake of completeness, I should mention two caveats. The first is that Fodor restricted his modules mainly to input-output cognitive mechanisms (such as perceptual and sensorimotor capacities), whereas subsequent theorists have posited that modules are more prevalent and may include a range of more central cognitive systems. The second concerns the precise nature of the connection between modularity and domain specificity. Fodor proposed the nine features of modularity in the spirit of necessary and sufficient conditions, but it may be more plausible to regard them as a looser cluster of features that are usually associated with modularity, though perhaps no single one of them is necessary.

mainly learned, such as chess-playing ability.[3] Having said that, we will have to revisit this link between domain specificity and innateness in the next section, after I have provided what I take to be the most defensible account of domain specificity in cognitive science. It will turn out that on that account, domain specificity only applies to innate cognitive kinds.

As for the link between domain specificity and *brain localization*, this is also widely made, as is the link between brain localization and modularity (which, as seen above, subsumes domain specificity). However, there does not seem to be a cogent reason for making either link. For instance, there are good grounds for thinking that various psychological capacities are modular (and hence domain-specific) even though they are not localized in one region of the brain, indeed even though they are scattered across a range of brain regions (vision and language are obvious examples). Modularity and domain specificity pertain largely to the functioning of a cognitive capacity rather than its neural implementation, so there is limited scope for inferring brain localization from either of these phenomena. Perhaps part of the reason for a conflation of the two notions is that the adjective "domain-specific" is sometimes loosely applied to a cognitive function *C* when it has been demonstrated that there is a *specific* neural mechanism *N* or a *specific* brain region *R* that subserves *C* (which does not subserve any other cognitive function, *C**). But there are other, more standard ways of referring to such a relationship, namely *selectivity*, or by saying that *N* is the *neural correlate* (or *neural substrate*) of *C*. Hence, use of the term "domain specificity" would appear to be misplaced in this context, and it is an unfortunate coincidence that the term "specificity" is sometimes used to denote neural specialization.[4]

In distinguishing domain specificity from other features of human cognition, I have so far relied on an implicit preliminary understanding of the phenomenon. As the term implies, what it is for a cognitive capacity to be domain-specific is for it to pertain to a single domain or to a restricted range of domains, and more importantly, for it not to be generalizable to

[3] In Khalidi (2001), I argued that when it comes to domain-specific abilities, it is easier to tell whether and to what extent they are innate or not. That is because we can more easily gauge the amount of explicit learning or relevant experience in the case of domain-specific cognitive capacities than in the case of domain-general ones. It is simply easier to rule out relevant sources of information in the former case than in the latter. But this does not mean that domain-specific abilities are more likely to be innate, just that the evidence is easier to assess.

[4] Here is a recent instance of this conflation from a paper in social neuroscience: "The brain is certainly not equipotential. However, there remain a number of interesting and difficult questions about the degree of such apparent *specialization*, how it might come about and what it accomplishes. These issues have been the focus of numerous theories of *domain specificity*, which range from abstract cognitive hypotheses to neurophysiological and neuroanatomical accounts" (Spunt & Adolphs 2017, 559; emphasis added). The purported link between domain specificity and neural localization will be revisited in Section 4.4.

other domains. Moreover, this last proviso highlights the importance of reserving domain specificity for aspects of cognition that are in principle *generalizable* across domains, although they are not in fact *generalized*. The idea is that a cognitive construct should be generalizable in principle, in the sense that it should be the kind of thing that could be generalized. It would be vacuous to describe as domain-specific some cognitive entity that is not even in principle generalizable. For example, a body of information pertaining to some domain or another (e.g. a list of world capital cities, or the entries in an address book) is not generalizable, since its subject matter is in principle restricted to a certain domain.[5] By contrast, a rule that is deployed by a cognizer in one domain but that *could* be deployed in another domain is in principle generalizable (e.g. modus ponens). Hence, we will mainly be concerned with domain-specific rules, principles, or algorithmic processes. These are the kinds of cognitive entities that can be generalized across domains and that are therefore candidates for being domain-specific, though they may not be the only cognitive capacities that can be so generalized.[6] Derivatively, cognitive capacities or systems can be characterized as being domain-specific when they comprise or include one or more domain-specific rules or processes. In the rest of this chapter, I will be speaking mainly of domain-specific capacities, but also sometimes, of rules, principles, or (algorithmic) processes. The next section will look at a particular example of a domain-specific capacity and use it to advance a proposal as to how to characterize domain specificity.

4.3 A Preliminary Example and a Theoretical Proposal

It will be helpful to begin with a relatively uncontroversial example of domain specificity, though I will argue that it is somewhat misleadingly described and improperly contrasted with a putative case of domain generality. Cosmides and Tooby (1994) explicate the well-known example of the alarm calls of vervet monkeys, who give three different calls in response to three different kinds of predators (leopard, eagle, and snake), leading conspecifics to take three different types of evasive action (respectively,

[5] There may seem to be a fairly simple sense in which one body of information is more generalizable than another, for example, a database containing all the residents of New York state, as compared to one containing all the residents of New York state. But this does not mean that the first database is more generalizable, just that it is more general or comprehensive.

[6] Compare Barrett (2018, 4): "One can then define the domain specificity of a process as the degree to which its operations vary across domains. A perfectly domain-specific process is one that operates only in a single domain and in no others. A perfectly domain-general process is one that operates identically across all domains."

https://doi.org/10.1017/9781009223645.005 Published online by Cambridge University Press

climbing a tree, looking up or diving into bushes, and standing on hind legs and looking into the grass) (cf. Cheney & Seyfarth 1990 cited in Chapter 1). In this case, they state: "A single, general-purpose alarm call (and response system) would be less effective because the recipients of the call would not know which of the three different and incompatible evasive actions to take" (Cosmides & Tooby 1994, 89–90). The problem with this observation is not that there could not be a general-purpose alarm system; there clearly could. But a general-purpose alarm system is *not* one that would issue the same call for every predator. That would be a system that fails to discriminate among different stimuli. Rather, an all-purpose alarm system would be one akin to a human linguistic alarm system, which issues a different linguistic warning in the case of different predators. There would clearly be certain advantages to such a system, since it would be capable of handling a much wider range of predators or dangers (e.g. "Lion!," "Hawk!," "Stampede of elephants!") and of being made more precise in various ways (e.g. "Tiger to the right," "Eagle to the northwest," "Human with weapon right behind you"). However, it may also involve certain disadvantages, since given the diversity of inputs and outputs, it may take more processing time to issue the correct alarm, there may be more opportunity for error in both transmission and reception, and the evasive action involved may have to be figured out from scratch by the respondent once the alarm is sounded. Determining which of these two alarm systems, the domain-specific vervet system or the domain-general human system, is more efficient and adaptive is not an easy matter. It will clearly depend on various contingencies such as the types of predators typically encountered, the seriousness and urgency of the threats they pose, and other features of the environment. Cosmides and Tooby may ultimately be right that in certain circumstances a domain-specific system may be superior to a domain-general one. But they do not appear to have drawn the distinction between domain specificity and domain generality in the right way.

This example is instructive since, once it is modified in the way that I have just done, it seems to provide a fairly clear contrast between a domain-specific and a domain-general cognitive capacity. The first feature that can be gleaned from the vervet monkey alarm call system is that some cognitive systems for alarm calls are at least in principle generalizable. That is to say, even though the vervet alarms are only issued for a small set of specific predators, it is not hard to conceive of a different alarm system that would extend to other predators (or indeed, to other types of stimuli). Hence, it seems safe to conclude that for one to speak meaningfully of a domain-specific cognitive capacity, it must have the following feature:

> (DS1) A *domain-specific* cognitive capacity is one that is in principle generalizable to new domains.

This condition may appear vacuous, but it is designed to rule out cognitive capacities that consist of a body of information or database rather than rules or algorithmic processes, as mentioned in the previous section. Domain specificity, to be meaningful, must be a feature of a cognitive capacity that is at least potentially generalizable. Though this point may seem obvious, the attribute of domain specificity is often conferred on bodies of knowledge or sets of concepts possessed by subjects that are not obviously generalizable, such as knowledge of animals as opposed to artifacts. The question of domain specificity would seem to be at issue only if those concepts are implicated in one or more inferential rules or principles, as when a set of concepts is embedded in a broader theory (along the lines of the theory theory of concepts encountered in Chapter 2). For example, there is evidence that animals are conceived of as having causal "essences," whereas artifacts are conceived of in terms of their function (Gottfried & Gelman 2005; Greif, Kemler Nelson, Keil, et al. 2006; cf. Boyer & Barrett 2005, 102). If so, then the inferential rules or principles associated with each theory might either be domain-general or -specific, depending on whether they are generalizable or not. Otherwise, it is not clear what sense to attach to the claim of domain specificity (or generality).

The first proposed feature of domain specificity makes reference to "new domains," which is a notion that is in need of further explication and justification. A new domain need not be what we might regard as an entirely disparate area of inquiry. Indeed, the underlying problem is that there are no ready-made boundaries that could serve to delimit domains. In the context of a discussion of whether human creativity is domain-general or domain-specific, Sternberg (2009, 25) poses the problem quite compellingly:

> The greatest challenge in understanding the domain generality versus specificity of creativity is in understanding the concept of a domain itself. Is literature a domain, or German literature, or modern German literature, or modern German literature in its original language, or what? Is cognitive psychology a domain, or psychology, or behavioral science, or social science? Because no consensual definition of a domain currently exists, it is impossible at this time to have a clear sense of exactly what domain-specificity means.

In fact, it is not clear that it will ever be possible to come up with a definition of a domain in the context of cognitive science, since knowledge or information does not come neatly divided into delimited parcels. In the case of the vervets, the original domain is thought to be something like: predators commonly encountered by vervets in the wild. An alarm system could *in principle* be generalizable to include the new domain: all predators, or even, all threats. The vervet alarm system is domain-specific because it appears to fail to generalize to these new stimuli. But these new stimuli do not, strictly speaking, have to be drawn from what we would normally consider to be another domain, such as a new sensory modality or a new area of inquiry. At this point, it might be asked, by virtue of what are they to be considered stimuli pertaining to a genuinely new domain? They must at least be stimuli that the cognizer has not encountered before. But that condition is surely too weak, since the domain-specific vervet alarm system clearly generalizes to new exemplars of leopards, eagles, and snakes, which the individual has not encountered before, indeed ones which perhaps no vervet monkey has encountered before. Rather, in this context, a plausible understanding of new stimuli is that they are ones that the system was not originally designed to cope with. This is admittedly a vague formulation and brings in thorny evolutionary considerations concerning the original design or *proper function* of an evolved trait (Millikan 1989; Neander 1991). Though it is not always easy to determine what the proper function of a cognitive capacity is, some reference to it seems inevitable, since cognitive capacities have evolved to fulfill a certain function. Accordingly, their generalizability consists in being able to extend beyond that original function to encompass cases that they were not designed to cope with, or ones that are not normally encountered in the environment in which they evolved. A similar conclusion has been reached by Boyer and Barrett (2005, 98), who write: "The domain of operation of the system is best circumscribed by evolutionary considerations." Barrett (2018, 6) expands on this conception, as follows:

> The actual domain of a computational process is the set of inputs for which the process can or will produce outputs whereas the proper domain is the set of inputs for which the process has been selected to produce outputs. This is roughly akin to the distinction between "selected effect" and "causal role" functions in biology: the selected effect function of a biological trait is the function it was selected to carry out (in the past) whereas the causal role function of a trait includes all of the functions it can in fact perform, whether or not it has been selected to do so …

On this evolutionary conception of domain specificity, a domain-specific capacity or process is one whose actual domain coincides with its proper domain, whereas a domain-general one is one whose actual domain outstrips its proper domain.[7] Not only does the evolutionary understanding of domain specificity allow us to delimit the boundaries of domains in a principled way, I will go on to argue in the rest of this section and the next that it enables us to make sense of the way in which the category of *domain specificity* is deployed in various areas of cognitive science, and results in a good candidate for a cognitive kind.

Based on the considerations just vetted, I propose that the second crucial feature of a domain-specific cognitive capacity is as follows:

> (DS2) A *domain-specific* cognitive capacity is one that systematically fails to yield a correct output, or fails to yield an output at all, in the case of inputs that the capacity did not evolve to deal with.

The need for this second feature can be further justified by reflecting on appropriate examples from the literature. One such case is provided in a classic paper by Cheney and Seyfarth (1985, 197), who describe the domain specificity of certain cognitive capacities in vervet monkeys, as follows:

> Within the social group, the behavior of monkeys suggests an understanding of causality, transitive inference, and the notion of reciprocity. Despite frequent opportunity and often strong selective pressure, however, comparable behavior does not readily emerge in dealings with other animal species or with inanimate objects.

In this example, both features outlined above are clearly in evidence. First, the principle of causality and the rule of transitivity are clearly applicable outside the realm of social interaction with conspecifics. The transitivity rule can be used to infer hierarchy relations among vervet monkeys (e.g. if *A* ranks higher than *B* and *B* ranks higher than *C*, then *A* ranks higher than *C*) but it can also be used to infer information about size, quantity, and other matters (e.g. if object *A* is larger than object *B* and *B* is larger than *C*, then *A* is larger than *C*). However, despite the clear applicability of this rule to domains that go beyond social interactions with conspecifics, Cheney and Seyfarth claim that vervets do not so apply the rule. Second, it is thought that vervets do not use the rule of transitivity on other species or inanimate objects simply because they evolved the rule to deal with the restricted domain of social interaction with conspecifics, which may have

[7] For a precursor to this construal of domain specificity, see also Sperber (1994) on the "actual domain" as opposed to the "proper domain" of a module. Sterelny (2003, 190) also delimits domains in terms of adaptation: "Domains correspond to related sets of adaptive problems environments pose for agents; problems which must be solved if the agent is to survive and reproduce." But he does not appear to explicitly characterize domain specificity and generality in terms of proper domains and actual domains.

been a more pressing adaptive problem. In this case, it may seem obvious that interactions with other animal species and with inanimate objects constitute genuinely new domains. There may not appear to be a need to use the second feature of domain specificity to justify the judgment that it does not generalize to genuinely new domains. But it is not a given that social hierarchy does not constitute a domain that also comprises other species, and that any thinker who could apply such a rule to one's own conspecifics could also apply it to members of other species. Hence, (DS2) can be used to confirm that this is indeed a case of domain specificity, since the actual domain of the cognitive capacity does not exceed its proper domain, which in both cases is social interactions with conspecifics.

4.4 Further Evidence

So far, the examples considered derive primarily from cognitive ethology, specifically studies on other primates. But the concept of domain specificity has also had considerable influence in cognitive neuroscience and developmental psychology. I will now consider whether the notion as I have characterized it can be pressed into service in other areas of cognitive science, particularly when it comes to humans.

There is a well-established body of evidence indicating the existence of category-specific semantic deficits in a range of patients with brain lesions and other neural abnormalities. However, the correct interpretation of this evidence remains a source of contention. Caramazza and colleagues have interpreted this evidence as indicating that semantic information is "domain-specific" (Caramazza & Shelton 1998; Caramazza & Mahon 2003). Other researchers have adopted different models to explain some of the same findings. Tyler and Moss (2001) hold that the selective deficits are an emergent phenomenon. Even though concepts are represented in a unitary distributed system, different types of concepts are structured differently. Since concepts in different domains have different internal structures, impairment of brain function leads to their being differentially affected (Tyler & Moss 2001). On the face of it, much of this evidence, and the surrounding debate, seems to pertain not to the question of domain specificity but rather to that of brain localization. When damage to a certain part of the brain results in selective impairment in naming animals but not plants or body parts, the question is whether this is evidence that representations underlying our semantic information concerning animals is localized in a particular area of the brain, or whether they are not localized but that some of them are more impaired than others by such damage.

Although this is an important question in its own right, it does not bear directly on domain specificity as such, as I have already argued (see Section 4.2). Similarly, neuroimaging data that has been brought to bear on this controversy is more pertinent to the question of localization rather than domain specificity. Caramazza and Mahon (2003, 358) think that "there clearly does seem to be neural differentiation by semantic category" based on neuroimaging data. But Tyler and Moss (2001, 246) find that: "The most striking aspect of the neuroimaging data is the extent to which living and non-living concepts activate common regions with only small and inconsistent differences between domains." The neuroimaging data is obtained mainly by testing healthy subjects on a variety of tasks (e.g. silent naming, word-picture matching) and then using various techniques (fMRI, PET scans) to determine whether different areas of the brain are differentially involved when processing content derived from different domains (e.g. animals, tools, food items). But this does not seem to enable us to draw conclusions regarding whether our capacity to think about such domains involves abilities that are generalizable or not. If our knowledge of animals activates different brain areas than our knowledge of tools, that does not mean that any cognitive abilities that range over such domains are restricted to these domains and cannot be applied to others.

At this point, it may be objected that there is at least an indirect connection between domain specificity and brain localization. If domain specificity is understood in evolutionary terms, as proposed in the previous section, then domain-specific cognitive capacities are adapted for a certain function. Given the dependence of cognition on the brain, it is likely that neural hardware would have been dedicated to carry out this function, which implies a certain degree of localization. Hence, the evidence from lesion patients and from neuroimaging studies bears on the question of domain specificity in the sense that localized capacities are more likely to be domain-specific (and domain-specific capacities are more likely to be localized). If we find that different regions are recruited in tasks involving, say, animals and artifacts, then it is likely that the capacity to reason about animals is indeed domain-specific, as is the capacity to reason about artifacts. But this inference is not warranted, for a couple of reasons. First, even though I have argued that all domain-specific capacities are adaptive, it is not the case that all adaptive capacities are domain-specific. Hence, adaptive cognitive capacities that receive dedicated neural resources are not necessarily domain-specific. Second, evolution exploits existing resources, and there is not likely to be dedicated neural hardware for every evolved cognitive function, as emphasized by recent work on neural reuse (see

Chapter 1 and references therein). The localization of certain cognitive capacities does not seem integrally linked to the question of whether they are restricted in their application, which is the notion of domain specificity being explicated in this chapter.

Among developmental psychologists, there are some longstanding debates concerning the domain specificity of our cognitive capacities. For example, as seen in earlier chapters, Carey and Spelke have argued that children have innate systems of knowledge that apply to distinct sets of entities and phenomena. Moreover, they add that the domains of human knowledge, such as knowledge of language, physical objects, and number, center on distinct principles. These "core principles" serve to distinguish one domain from another. But despite the fact that Carey and Spelke hold that our cognitive makeup consists of distinct domains, they also claim that conceptual change in these domains occurs in part by constructing mappings between these domains. For instance, mappings between the domains of physics and number play a role in children's reconceptualization of matter and material objects. Though the mapping is slow and difficult, children eventually succeed in using this mapping from one domain to another to differentiate the concept of weight from the concept of density (Carey & Spelke 1994, 191–192). But if one can transplant certain inferential principles from one domain to another, then those principles are likely not domain-specific. As I have already mentioned, it would be misguided to argue that a cognitive capacity is domain-specific merely on the grounds that it pertains to a distinct body of knowledge. The issue of domain specificity does not arise in such cases. Rather, generalizability of rules or principles is key, and in this instance that condition would seem to be satisfied, thus casting doubt on whether the capacities in question are truly domain-specific (though they may still be innate).[8]

Opponents of the claim of domain specificity also sometimes seem to aim their criticism at a different target. Bates (1994/2001) is at pains to distinguish the claim of domain specificity from claims of innateness and (brain) localization. She stresses that a cognitive capacity can have any two of these features without the third. However, her characterization of

[8] Here, it should be noted that domain specificity (like innateness, as explicated in Chapter 3) may be a matter of degree. In fact, it might be a matter of degree on (at least) three dimensions. These dimensions can be captured by the following questions: (1) How difficult is it to learn to apply the capacity beyond the proper domain to the actual domain (e.g. in terms of the amount of time required, or number of exemplars)? (2) How restricted is the scope of the actual domain beyond the proper domain? (3) What is the error rate in the actual domain as compared with the proper domain? These would seem to be the main ways in which we might try to quantify degrees of domain specificity.

domain specificity is vague; with respect to language, Bates (1994/2001, 134) says that the claim of domain specificity is that "localized language abilities are discontinuous from the rest of mind, separate and 'special' …" Moreover, despite her explicit cautionary notes, in presenting the arguments for and against domain specificity, she sometimes implicates innateness or brain localization instead. For example, she argues against the domain specificity of language on the grounds that the brain systems that support language show an extraordinary degree of neural plasticity (Bates 1994/2001, 139). She also characterizes the controversy over the domain specificity of language as follows: "Have we evolved new neural tissue, a new region or a special form of computation that deals with language, and language alone?" (Bates 1994/2001, 138) But, as already stated, that question does not have a direct bearing on the issue of whether knowledge of language or the capacity to learn language can be generalized to other domains. Whether or not there is a brain region that has evolved to deal with language alone concerns innateness and brain localization rather than domain specificity.

Another case for testing this account of domain specificity can be drawn from research on face recognition, perhaps one of the most widely discussed cognitive capacities in this regard. Researchers tend to be divided as to whether the human capacity to recognize the faces of conspecifics is a domain-specific capacity, or whether it is a capacity that is acquired as a result of more general cognitive processes, of the type used to acquire expertise in other areas of human cognition. Without trying to rehearse the voluminous evidence involved, I will mention just two findings that are pertinent to the issue of domain specificity. Humans do not develop expertise for recognizing the hands or bodies of conspecifics that is at all comparable to their expertise for recognizing their faces, as measured by accuracy and reaction time (McKone, Kanwisher, & Duchaine 2007, 12). Similarly, humans show decrease in accuracy in identifying faces when those faces are inverted but do not show such a decrease in identifying the facades of houses in the inversion condition (Yovel & Kanwisher 2004). The capacity to recognize upright faces rapidly and accurately does not seem to generalize to other visual stimuli. Object recognition is a skill that is in principle generalizable to domains beyond faces (e.g. hands, bodies, houses), but it fails to be so generalizable in humans. This is clearly in keeping with (DS1).[9]

[9] But see Boyer and Barrett (2005), who review some of evidence against the domain specificity of facial perception.

What of (DS2)? Though it is not always explicitly mentioned by researchers who work in this area, I propose that the evolutionary clause is at least implicitly assumed. Consider the following scenario. Suppose it were found that humans could indeed generalize their face recognition capacities to encompass the faces of dogs. Would this show decisively that the capacity is domain-general, after all? Proponents of domain specificity might not give up on their claim that this capacity is domain-specific, insisting that it is a domain-specific capacity that is specific to the domain of faces in general, or perhaps mammalian faces. Indeed, even if further evidence came to light suggesting that this capacity extends to other objects like the facades of houses, they might continue to posit that it is a domain-specific capacity dedicated to the detection of objects with certain salient parts in particular configurations. What would rule out such a challenge? As I argued earlier, there are no ready-made domains that would enable us to dismiss it in principle. Rather, it seems natural to say that such hypothetical data would not be evidence of domain specificity (across a broader domain) because of *evolutionary considerations*. Since it is likely that such a cognitive ability would have evolved to detect human faces rather than, say, faces of humans and dogs (given the relative recency of the domestication of dogs[10]), let alone the facades of houses, any extension beyond the domain of human faces is indeed a generalization of this ability, and an indication that it is not truly domain-specific. In fact, this is explicitly acknowledged by proponents of domain specificity in this area of research. McKone, Kanwisher, and Duchaine (2007, 12) hold that the domain-specific theory "proposes that a face template has developed through evolutionary processes, reflecting the extreme social importance of faces."

A final case often discussed in this regard is one already mentioned in the previous section, namely the finding that human thinkers, starting from an early age, make different categorization decisions and inferences when it comes to living things and artifacts. As mentioned in Chapter 2, even preschool children make a distinction between animals and artifacts, asking different types of questions when confronted with unfamiliar exemplars from each category (Greif, Kemler Nelson, Keil, et al. 2006). Many studies claim that children are "psychological essentialists" when it comes to living things, prioritizing internal properties and hidden causes, but not when it comes to artifacts, where they tend to rely on functions (e.g. Gelman 2004; Gottfried & Gelman 2005). This is usually interpreted as indicating that

[10] There is ongoing controversy about the dating of the domestication of dogs, but upper estimates suggest that it occurred less than 30,000 years ago.

they have domain-specific inferential rules that are applied to the domains of living things and artifacts, respectively (see Sloman, Lombrozo, & Malt 2007). In this case, the relevant rules are potentially extendable beyond each domain, yet are usually not so extended. For instance, though children could make inferences about artifacts based on their internal parts, or about living things based on their function, they generally refrain from doing so. Here, it may be thought that domains are well-defined and there is no need to cite evolutionary considerations in determining that these rules are indeed domain-specific. It seems that we may not need to lean on natural selection to make the claim that human thinkers have domain-specific capacities to reason about living things and artifacts, respectively. However, it is not so obvious when one considers evidence that does not fit neatly with this picture. For example, Keil (1996) suggests that the inferential principles associated with living things may sometimes be applied to complex artifacts, such as televisions, cars, and computers. Does this mean that they are domain-general after all? Not necessarily. Evolutionary considerations suggest that what may happen in these cases is that certain outward features of these artifacts (e.g. movement, interactivity) may "trick" the cognitive system into categorizing them as living things, at least in some contexts, and into applying the essentialist principles to them. Moreover, in many such cases, this leads to incorrect inferences concerning these artifacts, which means that systematic errors are made regarding them. That is why these principles remain domain-specific even though they may occasionally be extended beyond the boundaries of the domain of living things. It is not that the actual function of these essentialist principles exceeds their proper function, but that they rely on certain outward features to perform their proper function, effectively leading them to miscategorize some instances and make incorrect inferences. This analysis is somewhat speculative, but what matters is not that it is right but that questions of domain specificity or generality seem to be resolvable only against a background of evolutionary considerations.

Having tested the proposed characterization of domain specificity on a few examples, we are in a better position to briefly assess domain specificity as a cognitive kind. In keeping with the approach of this book, the status of *domain specificity* as a cognitive kind depends on its causal profile, both synchronic and diachronic. Focusing, first, on the etiology of domain-specific capacities, as characterized here, they are all such that they have been selected for some cognitive task and are restricted in their application beyond that task. That is, their actual function does not exceed their proper function. This means that they are individuated in part on the basis of their phylogenetic history. For this reason, Sperber (1994, 51) points out: "The domain of a module is … not a property of its internal structure (whether described in neurological or in computational terms)." Second, given that domain-specific capacities do not transcend their proper domains, this means that they lack a certain type of flexibility as compared with domain-general capacities. This is a synchronic causal property of domain-specific capacities and explains why they restrict cognizers in certain ways and do not equip them to deal with novel contexts or environments. Third, domain-specific systems may exhibit a certain kind of efficiency or superior ability when it comes to the particular type of cognitive task that they were selected for. By contrast, domain-general systems, which tend to be more flexible, are also less efficient. As Boyer and Barrett (2005, 100) observe: "In general, the more an inference system exploits external sources of information and stable aspects of the cognitive environments, the more computational power is required to home in on that information and derive inferences from it." This helps explain why domain-specific capacities tend to be fast and automatic (two other features associated with modularity): Since they are dedicated to a specific function and cannot outstrip it, the input–output mappings are likely fixed (which promotes speed) and preprogrammed (which makes them automatic). Hence, this causal profile may also help explain why some of the properties associated with modularity cluster together. This thumbnail sketch of the etiological and causal profile of domain-specific (and domain-general) cognitive capacities brings out some of the features that each class has in common and provides some reasons for thinking that they are genuine cognitive kinds.[11]

4.5 A Theoretical Challenge and Response

The chief theoretical challenge to an evolutionary account of domain specificity has been discussed in depth by Atkinson and Wheeler (2004). On their view, what they call the "grain problem," which is closely related to the problem of delimiting domains encountered above, persists even if we adopt an evolutionary understanding of domain specificity. Moreover, they think that the grain problem for domain specificity is just a special

[11] It should go without saying that proper domains need not be "philosophically correct" or "given by reality," as emphasized by Boyer and Barrett (2005). That is, they might not correspond to ontological divisions made by philosophers or scientists. "Faces are distinct objects only to an organism equipped with a special system that pays attention to the top front surface of conspecifics as a source of person-specific information" (Boyer & Barrett 2005, 98).

https://doi.org/10.1017/9781009223645.005 Published online by Cambridge University Press

case of a more general problem with evolutionary explanations that has been noted by Sterelny and Griffiths (1999). Atkinson and Wheeler (2004, 161) formulate the problem as follows:

> Is choosing a mate a single adaptive problem, or is it a set of related problems, such as: choosing someone of the opposite sex, someone who shows good reproductive prospects, and someone who shows signs of being a good parent? Or at a yet finer-grained level of description, is the problem of choosing someone with good reproductive prospects a single problem or a set of related problems, such as choosing someone who is young, healthy, of high status, etc.?

While Sterelny and Griffiths think that this problem can be overcome by adverting to the "device" or mechanism that is of interest in the explanation, Atkinson and Wheeler explain that even devices and mechanisms are subject to a grain problem. For example, in attempting to explain the adaptive function of a leg, do we focus on the adaptive function of the entire leg, or the foot, or a toe, or a bone in the toe? Does each of these have a separate adaptive function, or is it only the foot as a whole that has an adaptive function in its own right? This is why they think the grain problem has two dimensions, involving both the grain of the adaptive problem and the grain of the adaptive solution. Ultimately, Atkinson and Wheeler think that evolutionary explanations of cognition can bring the grain problem under control, but I think that they misrepresent the nature of the solution. To explain why, I will first consider their way out, then say why I think the resources for dealing with the grain problem are already implicit in the account of domain specificity that I have provided.

Atkinson and Wheeler's answer to the grain problem relies on the idea that by moving back and forth between the adaptive problem and the mechanism that is the solution, cognitive scientists and evolutionary psychologists can arrive at a satisfactory answer as to the domain specificity or generality of a cognitive capacity. As they put it: "Ideally, there is a dynamic and mutually constraining relationship between attempts to infer architectural solutions from adaptive problems and attempts to infer adaptive problems from architectural solutions" (Atkinson and Wheeler 2004, 169). They illustrate their solution with reference to research on emotions. This work has shown that "distinct neural structures are disproportionately involved in the perception of fear, on the one hand, and of disgust on the other" (Atkinson and Wheeler 2004, 165). Based on evidence from patients with neural damage as well as healthy subjects, the amygdala has been found to be involved in recognizing fearful facial expressions but not in recognizing expressions of disgust, happiness, and surprise. Meanwhile, the basal ganglia and insula are recruited in the recognition of facial expressions of disgust, but not of fear. Hence, they claim, there is a double dissociation between mechanisms involved in the recognition of the emotions of fear and disgust.[12] Meanwhile, evolutionary psychologists have provided reasons to think that the perception of fear and disgust emotions evolved under different selection pressures and for different reasons. For instance, there is selective pressure for a specialized system that detects potential danger and threat, and one way of detecting this is by perceiving fear in the faces of conspecifics. Moving back and forth between the implementational level and ecological level, Atkinson and Wheeler think that the grain problem can be finessed. Though they do not say so explicitly, it appears that they are claiming that in this case, at least given the current evidence, there is little difficulty in saying that the capacities to recognize emotions of fear and disgust are domain-specific capacities and that there is no domain-general capacity in humans to recognize emotions in the facial expressions of conspecifics. Both the brain mechanisms and the evolutionary scenario point to the conclusion that these are distinct capacities that have evolved to deal with distinct emotions. First, the fact that the two recognitional capacities are subserved by different neural mechanisms points to a different evolutionary origin for them and hence to the fact that capacities were designed to deal with different emotions. Second, there is a plausible evolutionary story to be told about the need for a separate system to detect threats and to register alarm in the faces of others.

Even though I am generally sympathetic to the conclusion they reach, I will argue that there are two flaws with Atkinson and Wheeler's analysis. First, as I have already argued, the neural implementation of a capacity is indirectly relevant to the question of its domain specificity or generality. At best, they have provided some weak circumstantial evidence to the effect that there is no single (domain-general) capacity to recognize facial expressions of emotion in conspecifics. But it is quite possible that there is a single domain-general system to detect emotion in the faces of conspecifics that is not localized in a single brain region, as opposed to two separate domain-specific systems to detect fear and disgust. Second, the real key to this case lies in the plausibility of the adaptive scenario that Atkinson and Wheeler allude to, which posits a different evolutionary origin for the

[12] In fact, contrary to Atkinson and Wheeler's claim, there is substantial evidence against double dissociation in this case, some of which has appeared since their paper was published. For instance, Bonnet, Comte, Tatu, et al. (2015) discuss the role of the amygdala in the perception of positive emotions, and Sander, Grafman, and Zalla (2003) reject the functional specialization of the amygdala. (I am grateful here to Dylan Ludwig for guidance.)

capacities to recognize fear and disgust. If that scenario can be sustained, it would show that there is probably no domain-general capacity to recognize emotion in the facial expressions of other members of our species. But notice that here as in other cases, one could still pose the question (concerning each of these capacities) as to whether there is a single domain-specific capacity that pertains to each of these emotions. For instance, someone could ask whether disgust perception is a domain-specific capacity that pertains simply to disgust, or a capacity that evolved to perceive disgust toward foods, say, and then was generalized to perceive disgust toward a range of other objects (e.g. human actions). The answer to this question would rest on the plausibility of the evolutionary evidence that one could produce in favor of either of the two scenarios. Indeed, when it comes not to the *recognition* of disgust in others but to the *emotion* of disgust itself, some cognitive scientists have posited that the food evaluation and rejection system was later applied to the evaluation and rejection of social action, so that feelings of disgust toward human action were generalized from feelings of disgust toward food (see e.g. Haidt & Joseph 2007). If this evolutionary scenario is itself plausible, and if there is some evidence to suggest that the recognitional capacity coevolved with the emotion itself, then a similar account might be given of our ability to recognize disgust in others (namely, that it initially evolved to recognize disgust toward food and was later generalized). In short, I contend that Atkinson and Wheeler have misrepresented the means of resolving the grain problem in this and in other cases. In each case, what enables us to determine whether a cognitive capacity can be considered domain-specific is an account of its proper function, in other words, the adaptive purpose that it evolved to serve. In accordance with requirement (DS2), if the cognitive capacity enables the organism to carry out tasks that go beyond its proper function, as determined by the most likely evolutionary scenario, then it can be considered domain-general; otherwise, it is domain-specific.[13]

In the previous example, it may seem odd to speak of a domain-general capacity for recognizing expressions of disgust in the faces of others (on the assumption that it evolved to detect expressions of disgust toward food). It may be objected that such a capacity would surely still be domain-specific,

[13] Another shortcoming of Atkinson and Wheeler's account is that it considers domain specificity in principle to be an attribute of either information or rules (2004, 151). But I would argue that a body of information that is fully specified and does not contain open variables always pertains to a certain domain, and hence the possibility of generalizability to new domains does not even arise in this case, as required by (DS1). That is why the only candidates for domain specificity are rules, principles, and algorithmic processes.

being restricted to disgust, as opposed to other emotional expressions such as fear. But despite the fact that it may seem counter-intuitive to say that a disgust-detection capacity is domain-general, on the evolutionary scenario that I sketched out, it would indeed be one that has actually been generalized by human cognizers beyond its proper function. Though disgust may seem a restricted domain by comparison with emotion, domain specificity has to do with the cognitive ability to go beyond a certain given domain, not the inherent specificity of that domain itself (indeed, one cannot even make sense of such a notion, in the absence of antecedently delimited domains). Our pre-theoretical intuitions about this and other cases may need to be revised in light of what we discover about the evolutionary history of the cognitive capacities that we examine.

Given its reliance on evolutionary considerations in delimiting domains, the account of domain specificity that I am proposing clearly only applies to evolved cognitive capacities. Since evolved phenotypic features of the organism are at least in part innate, it has turned out on this account that only innate capacities can be domain-specific. Hence, this would provide an integral link between innateness and domain specificity, contrary to what was claimed in Section 4.2, above. If the argument to this point has been correct, it is a link that results from the fact that a definite sense cannot be attached to the notion of a domain-specific capacity unless it is an adaptation of some kind, and hence innate. This may seem like a shortcoming of the account, since at least prima facie it seemed conceivable for a cognitive capacity to be domain-specific yet not innate. But if that hunch is correct, it would require us to find some other way of explicating domain specificity that does not cite evolutionary considerations.[14]

This account of domain specificity in terms of proper function may raise worries about evolutionary explanations in general, especially when it comes to psychological traits. It is difficult enough to reconstruct evolutionary history when it comes to morphological traits that leave physical traces, so it may seem like a thankless task to try to do so in the realm of cognition. Determining whether a cognitive capacity is domain-specific might be thought to be empirically intractable. But there has been a large body of scientific work on adaptive accounts of psychological

[14] What about cognitive capacities that acquire functions via their learning histories or ontogenetic development? Can we say whether they are domain-specific or -general, even though they do not have evolved proper functions? Though I will not try to justify this here, it may be possible to generalize the notion of function that I am exploiting to characterize domain specificity to cover such cases, along the lines proposed by Garson (2019) in his "selected effects" account of functions. (I am grateful to an anonymous referee for pressing me on this point.)

capacities and there are diverse pieces of evidence that can be brought to bear in trying to determine the proper function of a cognitive capacity. Ereshefsky (2007; 2012) has made a strong case for considering "psychological categories as homologies" and has outlined various methodologies and strategies that evolutionary biologists, ethologists, and others have for identifying "behavioral homology." This requires reconstructing phylogeny, as well as ontogeny, in an effort to determine the precise lines of descent of certain psychological traits. In fact, Ereshefsky contrasts the homological approach to investigating psychological categories with a functionalist one, on the grounds that functionalist accounts in psychology tend to be ahistorical and privilege synchronic causal roles. But in this chapter and elsewhere in this book, I have interpreted functions at least partly in etiological terms, and that accords with the approach of at least some of the cognitive scientists working on domain specificity (e.g. Sperber 1994; Sterelny 2003; Boyer & Barrett 2005). It is also in keeping with Ereshefsky's suggestion of combining functionalism with a homological approach to psychological traits, by building on the work of biologists who have argued that adaptational approaches to behavior and psychology have neglected the phylogeny and ontogeny of psychological traits (Ereshefsky 2007, 660). This approach and the various methodological strategies used by scientists in ethology, comparative cognition, and other areas, vindicate the feasibility of tracking the etiology of cognitive capacities in an attempt to ascertain their proper functions. As Ereshefsky (2007, 671) puts it: "The lesson for studying psychological categories as homologies is that regardless of our access to the historical record, an array of phylogenetic methods exists for testing adaptational hypotheses." Exploiting these methodological strategies should help us avoid positing just-so stories or falling into the crude adaptationism that characterizes some work in evolutionary psychology.

4.6 Conclusion

In this chapter, I have tried to provide an analysis of domain specificity in cognition that enables us to make a theoretically useful distinction between domain-specific and domain-general cognitive systems. Drawing on examples from the literature, both genuine and spurious, I have tried to show that there are two features that make a cognitive capacity domain-specific. First, the cognitive capacity must be one that is in principle generalizable. Hence, it cannot be something like a body of information concerning a particular area, but something more like a rule or algorithmic process that has wider applicability.[15] Second, it must be a capacity that the subject is unable to apply to genuinely new cases, where new cases are ones of a type that this system was not originally evolved to deal with, or ones that are not within what has been termed the *proper function* of this cognitive system. This second condition is important in that it provides us with a principled way of delimiting domains, since these are not antecedently given. Although it is often difficult to determine what the proper function of a cognitive capacity is or what it has evolved to deal with, since that requires us to understand the adaptive pressures that led to its being selected in the ancestral environment, this seems no more difficult in principle than the determination of evolutionary functions concerning other phenotypic features. This is what enables us to rule whether a cognitive capacity is one that the agent can extend to genuinely novel domains or not. The distinction between domain specificity and domain generality matters because a central debate in contemporary cognitive science concerns the extent to which our cognitive capacities are domain-specific tools or whether they are domain-general problem-solving capacities. A resolution of this disagreement depends on a clear means of demarcating domain-specific from domain-general systems or capacities. Moreover, it has often been claimed that one of the main points of difference between human cognition and that of other animals is its domain-general nature. Again, this debate cannot be properly adjudicated unless we have a principled way of making the distinction.

As a cognitive kind, domain specificity is related to innateness in at least two ways. The first is that it is a second-order kind, applying primarily to cognitive capacities or other first-order cognitive kinds. It may even be construed as a *property* of cognitive capacities, rules, or principles, since as I have pointed out at various points in this book, the distinction between kinds and properties may not run very deep, at least in cognitive science. The second way in which it is related to innateness is more direct. As argued in this chapter, according to this characterization of domain specificity, the only candidates for domain-specific cognitive kinds are evolved kinds. They are the ones that have proper functions that have been selected by the process of natural selection. These cognitive kinds are therefore innate to the mind, along with other evolved mental kinds. This means

[15] The question may be raised whether this means that domain-specific capacities are pitched at the algorithmic, not computational level. Though these capacities will include algorithmic processes, a judgment of domain specificity does not need to ascertain the precise algorithms involved or identify their main steps. The point is to determine the net effect of the capacity and its function, which is very much something to be determined at the computational level.

that domain-specific kinds are individuated, at least in part, with reference to phylogeny. Like some of the other cognitive kinds discussed (and to be discussed) in this book, it is identified on the basis of its causal history. As such, it does not supervene entirely on the intrinsic properties of the individual thinker but rather on the thinker's causal history and the causal history of the species. Part of what it is for a cognitive capacity to be domain-specific is for it to have a certain phylogeny, so if (to conjure a far-fetched scenario) two cognitive capacities are intrinsically identical they may not both be domain-specific if they do not share the same phylogeny. That is not to say that it will be easy to tell whether a cognitive capacity has the right etiology for it to qualify as a cognitive kind, but a definitive determination entails reconstructing phylogenetic history.