Concept Nativism and Neural Plasticity

Stephen Laurence and Eric Margolis

Abstract. One of the most important recent developments in the study of concepts has been the resurgence of interest in nativist accounts of the human conceptual system. However, many theorists suppose that a key feature of neural organization—the brain's plasticity—undermines the nativist approach to concept acquisition. We argue that, on the contrary, not only does the brain's plasticity fail to undermine concept nativism, but a detailed examination of the neurological evidence actually provides powerful support *for* concept nativism.

1. Introduction[*]

One of the most important recent developments in the study of concepts has been the resurgence of interest in nativist accounts of the human conceptual system. However, many theorists suppose that a key feature of neural organization—the brain's plasticity—undermines the nativist approach to concept acquisition. In this chapter, we will argue that, while this view of the matter has an initial air of plausibility, it gets things exactly backwards. Not only does the brain's plasticity fail to undermine concept nativism, but a detailed examination of the neurological evidence actually provides powerful support *for* concept nativism, giving rise to what we call the *argument from neural wiring*. The brain is most definitely an organ that is altered by experience, but as we will see, the ways in which it is altered support a nativist perspective on cognitive architecture.

2. The Contemporary Empiricism-Nativism Debate

Nativist views of concepts trace back to Plato, Descartes, Leibniz and other philosophers who theorized about the existence of innate ideas. Partly for this reason, nativism about the conceptual system is sometimes characterized simply as the view that there *are* innate ideas or concepts, and empiricism as the view that the mind is initially a blank slate in that it has no

innate structure whatsoever. However, this way of distinguishing empiricism from nativism would be ill-advised.

First, to characterize empiricism as the view that the mind begins with no innate structure would have the unfortunate consequence that there aren't really any empiricists. It has long been recognized by all parties to the empiricism-nativism debate that a mind without any innate structure—a truly blank slate—wouldn't be capable of learning. There has to be *something* that accounts for why human beings come to know anything at all about the world around them. As Quine once noted, even ''the behaviorist is knowingly and cheerfully up to his neck in innate mechanisms'' (Quine 1969, p. 95-96).[1] Second, while it is true that nativists are more likely than empiricists to embrace innate concepts in addition to other types of innate psychological structures, focusing exclusively on whether concepts are innate or not doesn't do justice to mainstream views of the conceptual system. Empiricists may accept some innate concepts, and nativists may hold that what matters is not innate concepts per se, but rather the existence of a rich innate basis for acquiring concepts.

For these reasons, we think it best to characterize concept nativism directly in terms of what is at stake in the disagreement between empiricists and nativists. For contemporary theorists in philosophy and cognitive science, the disagreement revolves around the character of the innate psychological structures that underlie concept acquisition.

According to empiricist approaches, there are few if any innate concepts and concept acquisition is, by and large, governed by a small number of innate general-purpose cognitive systems being repeatedly engaged. Sometimes this point is put by saying that empiricists claim that concepts are largely acquired on the basis of experience and hence that the conceptual system is predominantly a product of learning. But the crucial fact here isn't that empiricists place a lot of weight on learning. (As we'll see in a moment, nativists do too.) Rather, what is characteristic of empiricism in the empiricism-nativism debate is its *distinctively empiricist approach to learning.* The empiricist view is that concept learning overwhelming traces back to general-purpose cognitive systems and that these provide the psychological underpinning for the many varied concepts that humans come to possess. For example, on a typical empiricist view, concepts related to agency and concepts related to number are both the product of the same kind of psychological processes embedded in the same general-purpose concept acquisition systems. The reason why agency representations form in the one case and numerical representations in the other is simply a reflection of the differing experiences of the learner resulting in the different input to these systems.

The nativist approach, in contrast, holds that innate concepts and/or innate special-purpose cognitive systems (of varying degrees of specialization) play a key role in conceptual development, alongside general-purpose cognitive systems. So what is characteristic of nativism in the empiricism-nativism debate is its *distinctively nativist approach to concept acquisition,* crucially involving a substantial number of innate concepts, numerous innate special-purpose systems involved in concept acquisition, or, most likely, some combination of the two. A nativist

---

[1] Some contemporary theorists who undoubtedly fall on the empiricist side of the empiricism-nativism divide have rejected the label "empiricism" because of its association with the view that the mind lacks innate structure. For example, in a discussion relating work in neuroscience to theories of conceptual development, Steven Quartz remarks, "I have avoided using the term empiricism, instead stating the strategy in terms of not being strongly innate. My reason for this lies in the common identification of empiricism with *Tabula Rasa* learning" (Quartz 2003, p. 34). Since we take there to be a substantive issue at stake between theorists like Quartz and concept nativists, a better characterization of empiricism is clearly needed.

view is perfectly at home with the claim that representations of agency might depend on psychological processes that reflect the operation of innate agency-specific concept acquisition systems, while representations of number depend on separate, innate number-specific concept acquisition systems. The reason why agency representations form in the one case and numerical representations in the other would then be due as much to the fact that they are governed by different innate special-purpose acquisition systems as it is a reflection of the differing input to these systems.

As frameworks for explaining concept acquisition, both nativism and empiricism come in differing strengths. A strong form of empiricism would claim that there are no innate concepts whatsoever and that concept acquisition depends exclusively on a small number of general-purpose psychological systems. Jesse Prinz defends a view along these lines, holding that concepts "are all learned, not innate" (Prinz 2005, p. 679). After arguing against what he takes to be the main proposals for special-purpose innate concept acquisition systems, he summarizes his discussion by noting, "I do not believe that any of these domains is innate. That is to say, I do not think we have innate domain-specific knowledge that contributes to structuring our concepts" (Prinz 2005, p. 688).

A weaker empiricist view might admit that there are a limited number of special-purpose cognitive mechanisms that constrain how the conceptual system develops in certain isolated cases, particularly cases supporting basic biological needs, but apart from these few minor exceptions, concept acquisition is governed solely by general-purpose cognitive systems. Rogers and McClelland (2004) defend such a view. After arguing that a general-purpose connectionist model can explain the way that adult semantic memory is organized, they suggest that there may be a handful of instances of prepared learning, including, for example, a tendency to withdraw from strong stimuli and to respond favorably to the taste of fat and sugar. Nonetheless, Rogers and McClelland state that they are "reluctant … to accept that, in general, human semantic cognition is prepared in this way", arguing instead that "domain-general mechanisms can discover the sorts of domain-specific principles that are evident in the behavior of young children" (Rogers and McClelland 2004, p. 369).

Nativist views also come in differing strengths. In fact, arguably one of the main reasons why concept nativism has had so few adherents is because of its association with one of the most audacious nativist positions ever defended—Jerry Fodor's radical concept nativism (Fodor 1975, 1981). According to this view, nearly all concepts corresponding to individual words in natural languages are innate, including the likes of LINGUINI, CARBURETOR, BEATNIK, and QUARK. Notice that it isn't just the sheer volume of innate concepts that makes this view so outrageous—the thousands and thousands of concepts corresponding to actual and potential natural language words—but also the fact that most of these concepts are clearly newcomers in human history, dependent upon specific historical cultural and technological conditions for their appearance. Fodor once tried to explain how it might be that these many innate concepts could be sitting around unused for much of human history only to suddenly become active. His suggestion was that they get "triggered" by innately specified environmental conditions whose occurrence might depend upon highly contingent prior events. But this really is a singularly implausible maneuver on his part, and Fodor's view has found few if any advocates even among nativists.[2] Fodor's

---

[2] Chomsky (1991) flirted with Fodor's radical concept nativism, but even Fodor himself has now rejected the view (Fodor 1998, 2008).

radical concept nativism is not just an outlier position within the empiricism-nativism debate, it is also an outlier position on the *nativist* side.[3]

What would a more reasonable concept nativist position look like? First, while denying that all or virtually all concepts are innate, it would nonetheless embrace the existence of a substantial number of innate concepts, including concepts that pick out abstract categories. Second, it would embrace a variety of innate special-purpose acquisition systems that are geared to particular conceptual domains. And third, it would also embrace concepts acquired via relatively general-purpose innate acquisition systems—indeed, any tenable form of concept nativism will acknowledge that a great deal of the conceptual system is acquired, at least in part, by such general-purpose systems.

Which concepts and special-purpose acquisition systems should a nativist say are innate? Ultimately, of course, this is a empirical question; there is no specific list that a concept nativist must be committed to. In our view, though, likely contenders include concepts and innate special-purpose systems associated with the representation of objects, physical causation, distance, movement, space, time, geometry, agency, goals, perception, emotions, thought, biological kinds, life stages, disease, tools, predators, prey, food, danger, sex, kinship, group membership, dominance, status, norms, morality, logic, and number. Note, however, that this isn't to advocate the grossly implausible claim that *all* concepts related to these domains are innate or acquired via innate special purpose systems. For example, any sensible form of concept nativism would reject the idea that PASTA and BONBON are innate, but might nonetheless accept that there is an innate special-purpose acquisition system, or set of systems, that is involved in the conceptualization of food. Likewise, concept nativists needn't suppose that the concepts for advanced mathematics are innate, but may well maintain that there are innate numerical representations of one kind or another.

Earlier we noted that it would be wrong to suppose that only empiricists hold that concepts are learned. We can now see why. This is because nativism isn't confined to postulating innate concepts. Rather, a big part of concept nativism is its appeal to innate special-purpose systems of acquisition, and these systems are often best understood as learning systems. For instance, an innate special-purpose system for food might support the learning of which items in the environment are to be eaten and which are to be avoided, guiding food preferences and food-seeking behavior. Or an innate special-purpose system for faces might support the learning of concepts of individuals. These hypothesized systems are very much in the business of learning about the world, according to the nativist. They are just specialized for learning particular information in a way that is highly constrained by the nature of the learning system.

To a large extent, then, the difference between nativism and empiricism isn't *whether* learning is central to human concept acquisition but rather their differing views of *how* learning works. While empiricists take learning to be almost exclusively mediated by innate general-purpose learning systems, nativists maintain that general-purpose learning systems, though real and important, are not sufficient, and also postulate numerous innate special-purpose learning systems, holding that they are central to conceptual development, to categorization, and to other concept-involving higher-level cognitive processes.

It is also worth emphasizing that nativists don't deny the existence of relatively domain-general concept acquisition, although nativists are likely to see even domain-general acquisition as typically relying upon special-purpose systems or constraints of one kind or another. For

---

[3] See, for example, Laurence and Margolis (2002) and Pinker (2007) for highly critical assessments of Fodor's radical concept nativism that are nonetheless quite congenial to nativism in general.

example, socially mediated learning undoubtedly facilitates the acquisition of wide range of concepts, including concepts that are peculiar to a learner's culture. But on a nativist view, this learning is likely to be mediated by innate systems for acquiring cultural norms (Sripada & Stich 2006), by innate systems that are responsive to pedagogical cues (Csibra & Gergely 2011), and by innate biases governing cultural transmission, including biases regarding who to imitate (Boyd & Richerson 1985).

In the rest of the chapter, we will take up the question of how the facts related to neural structure and function bear on the evaluation of concept nativism, but before we do that, we should briefly comment on the status of the argument we will be offering for concept nativism—the argument from neural wiring—and how it fits into the larger case for the nativist framework. This argument is intended as a nondemonstative argument that takes the form of an inference to the best explanation. And so while we maintain that the argument provides strong support for concept nativism, the support is defeasible. We take it that the status of concept nativism turns entirely on nondemonstrative arguments of this kind. Empiricists sometimes write as though empiricism is the default position and that nothing short of an incontestable proof for concept nativism should move us away from this default (see, e.g., Prinz 2005, 2012). However, this outlook is misplaced. Given that empiricists and nativists disagree about what is clearly a factual question about the structure of the mind, both sides are equally in need of evidence and argument to establish their view, and must equally make the case that the balance of empirical considerations stand in its favor. Proofs have no more place here than they do in other disputes about the workings of the mind. Any nondemonstrative argument for concept nativism will also need to be considered in light of the other empirical arguments that bear on the empiricism-nativism debate. Although we don't have the space in this chapter to examine other arguments, we believe that there are important arguments for concept nativism drawing on evidence from developmental psychology, animal psychology, evolutionary psychology, anthropology, and other fields, that these arguments are mutually reinforcing, and that it is this total package that makes concept nativism such an attractive view.[4] Fortunately, though, the argument from neural wiring offers considerable support for concept nativism even taken in isolation. So while we don't want to give the impression that concept nativism stands or falls with this one argument, we do think the argument from neural wiring constitutes a solid reason for favoring concept nativism all the same.

3. Plasticity As a Challenge to Nativism

As we noted above, plasticity is often seen as providing an argument against nativism, not for it. We will begin, then, with the considerations that are generally thought to show that neural plasticity poses a serious challenge to concept nativism.

Plasticity has become a catchall term for the many ways that neural organization and function change in response to an animal's experience and action, as well as to traumas to its brain and body. A standard example that is often used to illustrate the general idea is the reorganization of the sensory map of the hand in the brain in response to an injury or to changes in the way that the hand is used. The normal arrangement in the somatosensory cortex is for adjacent groups of cells to correspond to adjacent regions of the body—for example, neurons

---

[4] For detailed discussion of a broad range of arguments for concept nativism and their interrelations, see Laurence and Margolis (in prep.)

that respond to the index finger reside close to neurons that respond to the middle finger. What happens if the nerve fibers connecting a finger to the spinal cord are severed? One might expect the cortical area that previously responded to that finger to atrophy, but this is not what happens. Instead, this cortical area is taken over by the adjacent finger(s), changing the function of that cortical area so that it responds to the adjacent finger(s). Likewise, even if one of two adjacent fingers is simply used more or happens to receive a greater amount of stimulation, some of the neurons that were originally responsive to the less active or less stimulated finger become responsive to the more stimulated finger. As one major textbook summarizes the matter, "This functional plasticity suggests that the adult cortex is a dynamic place where changes can still happen. Such phenomena demonstrate a remarkable plasticity" (Gazzaniga, Irvy, and Mangun 2009, p.102).

Neural plasticity doesn't stop with potential changes to the boundaries of somatosensory representations. Even more interesting are instances in which cortical circuits deprived of their usual sensory input come to be recruited by a differing sensory modality. An important example of this type of reorganization can be found in early-blind Braille readers. When tested on *tactile* discrimination tasks, they show increased activation in the *visual* cortex compared to sighted subjects, who show deactivation in this area (Sadato et al. 1996).[5] Studies using repetitive transcranial magnetic stimulation (rTMS) further confirm the role of visual cortex in the blind. rTMS is a technique that employs a magnetic pulse to disrupt neural activity in a targeted region of the brain. When rTMS is used to disrupt neural activity in the occipital (visual) cortex, blind subjects have difficulty identifying Braille and embossed Roman letters, whereas there is no effect on sighted subjects engaged in comparable tactile discrimination tasks (Cohen et al. 1997).

Perhaps the most celebrated instance of neural plasticity comes from a study with ferrets in which ferrets' brains were surgically rewired shortly after birth (Sharma, Angelucci, and Sur 2000; von Melchner, Pallas, and Sur 2000). Retinal projections were rerouted so that the neural signals that would normally go to the primary visual cortex were fed to the primary auditory cortex via the auditory thalamus. When tested as adults, not only did the auditory cortex in ferrets with rewired brains come to exhibit patterns of activity that are characteristic of the visual cortex (e.g., it contained groups of cells that responded differentially to stimulus orientation), but the ferrets were able to approach objects that could only be detected by sight. The take-home message, according to one of the original research reports announcing these results, is that "the pattern of early sensory activation can instruct the functional architecture of cortex to a significant extent" (Sharma, Angelucci, and Sur 2000, p. 846).

There is, of course, a question of how much of the brain exhibits this level of plasticity. Is plasticity unique to sensory-perceptual systems? One reason to think it isn't is that there are cases in which children recover from focal damage to cortical areas involved in language. Amazingly, there are even cases of children who come to develop near-normal linguistic abilities after undergoing a hemispherectomy, in which one cerebral hemisphere is disabled or entirely removed (Curtiss and Schaeffer 2005, Curtiss and de Bode 2003). One child (known as EB) who underwent a left hemispherectomy at 2.5 years old managed to recover much of his language

---

[5] Early-blind subjects include those who are blind from birth as well as those who developed blindness early in childhood. In the study by Sadato et al., this includes children who were blind before the age of seven, while in some of the other studies we cite below, the cutoff point for being considered an early-blind subject is at significantly earlier ages.

skills two years later.[6] When tested at the age of 14, his language was found to be normal in most respects, with all of his linguistic functions now residing solely in his right hemisphere (Danelli et al. 2013).

There may also be general theoretical reasons to suppose that the parts of the brain that are involved in higher cognitive processes are plastic and hence can take on any number of differing functions. Buller and Hardcastle (2000) argue that this is extremely likely given that dedicated brain circuits per se don't exist even for "our most basic [i.e., sensory] processes" (p. 313). In general, they claim, "the dedication of a brain system to a particular task domain is subject to change as the inputs to that brain system change" (p. 313). The reason there is the appearance of isolable and stable neural structures is that "plastic human brains have encountered recurrent environmental demands throughout history" (p. 317). Elman et al. (1996) express a similar view. After reviewing the evidence regarding brain plasticity in animals and humans, they conclude that there may be some "primitive innate representations in the midbrain…, but the cortex appears to be an organ of plasticity, a self-organizing and experience-sensitive network of representations that emerge progressively across the course of development" (Elman et al. 1996, p. 315).

It is worth noting that singling out the cortex isn't uncommon among empiricists who place a lot of weight on brain plasticity. Quartz (2003), for example, disagrees with Elman et al. about subcortical structures, maintaining instead that innate neural systems are the norm in most animals. But he agrees that a significant level of plasticity is a peculiar feature of the human cortex, pointing, in particular, to the example of how the occipital cortex functions differently in blind Braille readers and sighted control subjects (as in the work by Sadato et al. 1996 described above). According to Quartz, "The sharp contrast between cortical and subcortical structures suggests that the evolution of cortex may represent the evolution of a new acquisition strategy" (Quartz 2003, p. 36).

In sum, there is a lot to be said for the view that the brain exhibits a great deal of plasticity in terms of its functional structure. It also isn't difficult to see why this fact is often thought to favor empiricist approaches to conceptual development. Widespread and significant instances of neural plasticity suggests an inherent openness to the functions that any cortical area can take on. If this is right, then the brain's concept acquisition capacities needn't be innately constrained toward any particular outcome. Instead, cortical circuits might simply form as required to accommodate a learner's needs given whatever contingent sensory input has been received and the wiring that has been previously established. Buller and Hardcastle capture this picture in suggesting that neural plasticity goes hand in hand with the idea of a content-neutral capacity for addressing information-processing tasks. As they put it, "Our ancestors may have encountered diverse problems, but the brain evolved a *general solution* to those problems" (Buller and Hardcastle 2000, p. 317; italics added).

4. The Argument from Neural Wiring

We have seen that the phenomenon of neural plasticity appears to constitute a major objection to concept nativism. The more equipotential the brain, the less plausible it is that there are innate special-purpose acquisition systems. And examples like the ones reviewed in the last section do

---

[6] In right-handed individuals, like EB before his surgery, language is usually controlled by structures in the left hemisphere.

seem to point in the direction of a highly equipotential brain with a functional organization that fits naturally with an empiricist approach to cognitive and conceptual development. However, appearances are misleading. In this section, we will argue that not only do the facts pertaining to neural plasticity fail to discredit concept nativism, but a careful look at the relevant neurological evidence strongly favors the nativist approach. What we are calling the *argument from neural wiring* is an argument for nativism that draws upon a broad range of neurological evidence showing that neural plasticity is highly constrained in ways that are best explained within a nativist framework.

4.1 Plasticity Revisited

We will begin with some observations about a few of the striking examples from the last section. It is instructive to see that even in these instances, which are often thought to make nativist views look hopeless, the changes that the brain undergoes aren't as flexible and open-ended as they first appear.

Consider the finding that language can be recovered after the loss of the cortical areas that normally support linguistic processing, even with the loss of as much as half of the cortex. We saw this with EB, who had his left cerebral hemisphere removed when he was just two years old and yet grew up to have near-normal language abilities. In his daily life, he doesn't exhibit any noticeable signs of language impairment. Clinical assessments have also found his language to be near normal (compared to age-matched controls), with below average performance in only a few tests (e.g., in reading he had difficulty with homophones and words that have been borrowed from other languages). These minor difficulties only serve to highlight how strong his core linguistic abilities are. EB is a particularly interesting case study among hemispherectomy patients because the pathology he suffered didn't involve epilepsy and was localized in one hemisphere, leaving the other hemisphere intact and healthy. The pathologies that lead to hemispherectomy aren't usually so localized; generally there is damage to both hemispheres.

At first glance, then, this may suggest that EB offers a clear-cut illustration of the way that the plasticity of the brain stands in opposition to a nativist theory of development. EB's right hemisphere took on cognitive functions that are usually located elsewhere in the brain—a massive relocation of the neural circuitry for language (and for much else). But functional magnetic imaging (fMRI) shows that the areas in his right hemisphere supporting his linguistic abilities aren't scattered in unpredictable ways, as one might expect if his recovery were based on the powers of a truly equipotential brain. Rather, EB's language areas are homologues of the left hemisphere areas activated in linguistic tasks in healthy control subjects.[7] Danielli et al.

---

[7] Contrary findings in clinical and neuroimaging studies have led to a controversy about whether there are language-specific neural circuits—circuits that are specialized for and selectively activated by language processing. On the one hand, clinical studies have found sharp dissociations between linguistic and nonlinguistic deficits, suggesting the existence of language-specific areas of the brain. But on the other hand, neuroimaging studies with neurotypical subjects have failed to show processing regions exclusively devoted to language processing. However, the problem with the neuroimaging work that fed this controversy is that the selection of the brain regions for examination didn't take into account individual differences in neuroanatomy. Once the classic language areas are singled out in a way that is sensitive to these differences, fMRI studies reveal the existence of neural regions selectively activated by linguistic processes to the exclusion of other cognitive processes (such as mental arithmetic and general working memory) in accordance with the clinical data (Fedorenko, Bher, and Kanwisher 2011).

(2013) performed an in-depth fMRI analysis of language production vs. comprehension, automatic vs. controlled language processes, and auditory vs. visual processing, finding that "the overall neurofunctional architecture of EB's right hemispheric language system mirrors a left-like linguistic neural blueprint" (p. 225). This result points to a highly constrained neural organization that illustrates an important general principle: the brain's two hemispheres incorporate a large measure of potential redundancy of function that can be exploited at certain stages of development. This is a form of plasticity, to be sure, but not a kind that favors empiricism.

What about the ferrets? Certainly this at least illustrates true equipotentiality? After all, their rewired brains led to the auditory cortex developing features of the visual cortex, and even to the ferrets being able to respond to visual stimuli relying on processing that could only occur in their auditory cortex. Surprisingly, even this case fails to provide a strong argument against nativist theories of conceptual and cognitive development.

For one thing, generalizing from one relatively small area of the auditory cortex to the rest of the brain is a huge leap (Pinker 2002). Even if it turns out that this portion of the auditory cortex and its downstream perceptual processing is highly malleable in a way that reflects a given input modality, it doesn't follow that the processing for higher cognitive functions can be handled by arbitrary cortical areas. Higher cognitive processes may well involve dedicated brain areas even if sensory processes do not.

In the previous section, we encountered an argument by Buller and Hardcastle suggesting the opposite conclusion. Their claim was that sensory processes are among the most "basic" in the cortex and consequently that if there are no dedicated brain areas for these processes, then it is even less likely that there would be dedicated areas for higher cognitive processes. However, a lot is being packed into the idea of a so-called basic psychological process in this argument. It may be true that, when the brain receives external environmental stimulation, sensory processes are typically activated earlier than higher cognitive processes. However, this isn't enough to get Buller and Hardcastle's argument off the ground. What they need is for sensory processes to be more fundamental than higher cognitive processes in ontogeny—for the development of the neural systems for cognitive processes to depend upon the prior development of, and input from, the neural systems for sensory processes. But this claim would be question-begging in the present context. Whether there is such a dependence is partly what empiricists and nativists disagree about when the debate turns to the wiring of the brain.

Moreover, the point about the danger of an overgeneralization isn't just about the divide between sensation and cognition. It also applies to the internal structure of perceptual systems as well. The impact of the rewiring experiments was that the primary auditory cortex in a ferret with a rewired brain was able to process information from the retina that normally would have been processed by the primary visual cortex, and consequently came to take on some of the features that are characteristic of the primary visual cortex. The primary visual cortex is connected to a large number of distinct brain regions that support further specific types of visual processing, including computations responsible for downstream representations of location, direction of motion, speed, shape, and so on. If the redirected sensory input in the ferrets had led to the development of all of this downstream structure in the auditory cortex, that would amount to exactly the sort equipotentiality of cortical areas that empiricists could use against nativists. But, in fact, none of this downstream structure was reproduced. The overall wiring of the ferrets' auditory cortex was largely unchanged (Majewska and Sur 2006).

This situation—in which the primary auditory cortex in some sense allowed the ferrets to see despite the fact that downstream auditory areas weren't significantly altered—leads to a puzzle. *How is it that the primary auditory cortex can process and make any use of visual information?* Interestingly, Sur, the team leader in the ferret study, has answered that "the animals with visual inputs induced into the auditory pathway provide a different window on some of the same operations that should occur normally in auditory thalamus and cortex" (Sur 1988, p. 45; quoted in Pinker 2002, p. 96). In other words, the processing that is supposed to be handled by these different sensory areas is somewhat similar. Pinker expands on this idea by noting that there are general high-level likenesses between the computations that might be expected to take place in hearing and vision, as soundmakers with different pitches may be treated like objects in different locations, and sharp changes in pitch may be treated like motions in space (Bregman and Pinker 1978). If this is right, then even though the rewiring experiments show that the auditory cortex can be recruited for a certain amount of visual processing, this is because the auditory cortex and the visual cortex overlap in the types of computations they naturally support. Far from being a model case of the environment instructing an equipotential cortex, Sur et al.'s rewiring experiments illustrate the way that cortical structure and function remains largely unchanged even in the extreme case of input coming from a different sensory system.

4.2 Constrained Plasticity in Neural Structural Organization

So far, our response to the appeal to plasticity as an argument against concept nativism has been to show that some of the flagship examples of plasticity don't really amount to the sort of flexibility of neural organization that would tell against the nativist approach to cognitive and conceptual development. We now turn to evidence related to the wiring of the brain that directly supports the nativist perspective. The general form of our argument will be to point to aspects of neural, cognitive, and conceptual development that exhibit *constrained plasticity*—development that is not open-endedly plastic, but instead is highly constrained in ways that suggest important innate biases, predispositions, and limits on neural structure and function. We begin, in this section, by briefly presenting some evidence indicating that neural structural organization is not primarily driven by environmental input or feedback configuring the brain.

In a landmark investigation, Verhage et al. (2000) examined neurological development in a group of mutant (or "knockout") mice whose brains (as a result of the genetic mutation) were unable to release any neurotransmitters and thus were deprived of all synaptic transmission.[8] Accordingly, these mice would have had no experience-driven neural development whatsoever. Verhage et al. (2000) compared these mice to control littermates, and found that up until birth their brains were remarkably similar.[9] As Verhage and colleagues explain:

> Despite the general, complete, and permanent loss of synaptic transmission in the knockout mice, their brains were assembled correctly (Fig. 3). Neuronal proliferation, migration, and differentiation into specific brain areas were unaffected. At E12 [embryonic day 12], brains from null mutant and control littermates were morphologically indistinguishable (Fig. 3, A and B) … . At birth, late-forming brain areas such as the neocortex appeared identical in null mutant and control littermates, including a distinctive segregation of neurons into cortical layers (Fig. 3, C and

---

[8] Specifically, Verhage et al. suppressed the expression of the *munc18-1* gene in the mutants.

[9] After birth, of course, the mutant mice died. Without functioning synaptic communication, the brain can't even support the most basic life functions, such as breathing.

D). Furthermore, fiber pathways were targeted correctly in null mutants… (Fig. 3, G and H). (Verhage et al. 2000, p. 866; figure references are to figure 3 in the original article, reproduced here as figure 1.)

[Figure 1 about here]

This degree of similarity shows that many features of even the fine-grained structure of the brain can develop without any sensory input or feedback. Experience-driven neural activity may play more of a role in fine-tuning and maintaining this layout, rather than in establishing the overall organization itself (Marcus 2004). None of this is to say that the brain isn't plastic to some degree.[10] But it underscores the fact that neural development isn't open-endedly plastic.

4.3 Constrained Plasticity in Functional Organization

Constrained plasticity isn't confined to neural structural organization. It also extends to the functional specificity associated with particular cortical areas. This is especially apparent in instances of congenital and early sensory deficits, such as blindness, in which the specific cognitive function of the affected cortical area is preserved. Preserved function in these cases argues for constrained plasticity because it shows that the information processing associated with the region isn't dictated by the sensory information it receives.

Viewed in a certain light, congenital sensory deficits are essentially rewiring experiments. Cortical areas that usually process sensory information from an impaired modality (e.g., subareas of the visual cortex in the congenitally blind) end up taking input from another sensory modality, an organization (or reorganization) known as *crossmodal plasticity*. An empiricist who supposes that the cortex has a high degree of equipotentiality would predict that the resulting functions of these cortical areas would differ from the normal case, reflecting the difference in the input. But recent work with congenitally blind subjects shows that, contrary to this prediction, what typically happens is that downstream components of the visual cortex and related brain areas have the same functional specificity in the congenitally blind as in sighted individuals.

Consider, for example, the representation of the spatial location of objects in the visual cortex. In one study, Renier et al. (2010) presented early-blind subjects with auditory stimuli that varied in terms of sound type (different piano chords) and spatial location.[11] Using fMRI, their brain activity was measured during two behavioral conditions, an identification condition (in which they had to determine whether sequentially presented stimuli were of the same sound type) and a location condition (in which they had to determine whether they had the same location). Renier et al. found that the anterior part of the right middle occipital gyrus (MOG)—a part of the visual cortex associated with the representation of visual spatial location in sighted subjects—was differentially active for the auditory spatial localization task relative to the auditory identification task. Renier et al. also ran an analogous tactile task with the same subjects. In this case, their fingertips were given different types of stimulation (for the

---

[10] Notice that it shouldn't be controversial that the brain is plastic in the minimal sense that changes to the brain occur as people learn, think, and experience the world. This is simply a requirement on any broadly materialist theory of the mind.

[11] In this experiment, the early-blind subjects were either blind from birth or by the second year of life, without ever having had normal vision or, at the time of testing, memories of visual experience.

identification task) or there was stimulation to different fingers (for the spatial location task). Once again, fMRI data revealed that the right anterior MOG was differentially active for the spatial localization task relative to the identification task.[12] The upshot of this study is that, while the MOG is clearly plastic—in the early blind it comes to subserve auditory and tactile spatial localization abilities that it does not subserve in sighted individuals—the plasticity it exhibits is a form of constrained plasticity. The MOG continues to carry out the function of spatial localization in the early blind, just with different types of sensory input.

Further fMRI studies have revealed the same pattern of constrained plasticity in other neural regions with associated spatial functions. For example, Lingnau et al. (2012) examined the activity in the posterior parietal cortex (PPC), which is involved in the representation of space for purposes of guiding action. The PPC normally takes its sensory input primarily from vision and exhibits a pronounced gradient—with posterior subregions recruited more heavily for (visual) guidance of reaching and grasping, and anterior subregions more for planning and execution of motor action. Using a proprioceptively-guided reaching task, however, Lingnau et al. were able to show that the same pattern of functional differentiation occurs in congenitally blind subjects. Lingnau et al. compared the brain activity of congenitally blind subjects and blindfolded sighted subjects performing one of two actions on an object (touching with fingertips vs. grasping with their whole hand) in a specified location (chosen from five possible locations). In this case, the principle of constrained plasticity predicts that sighted and blind subjects would have similar activation in the anterior portions of the PPC (since all subjects would be equally engaged in the planning and execution of motor action), but that there would be significantly greater activation in the posterior portions of the PPC in the blind (since in the blind, the PPC would be accommodated to *nonvisual* sources of information regarding spatial location). This is exactly what Lingnau et al. found, leading them to conclude that "neural plasticity acts within a relatively rigid framework of predetermined functional specialization" (Lingnau et al. 2012, p. 7).

One of the major features of the visual cortex is the functional division corresponding to two broad networks of interrelated neural regions. The ventral visual stream (the "what" pathway) represents object properties, and is involved in object recognition; the dorsal visual stream (the "where" pathway) represents object location and the spatial relations between objects, and is involved in object-directed action. The results from Renier et al. (2010) and Lingnau et al. (2012) that we have been reviewing indicate that the dorsal visual stream continues to exist in early-blind and congenitally blind subjects and that its component subregions engage in the same functional processing for object location despite profound changes in sensory input (auditory or proprioceptive vs. visual).

Further studies of the dorsal visual stream fill out this picture by showing that that it's not just the representation of spatial location that is preserved. For example, Wolbers, Zahorik, and Giudice (2011) examined activity in the dorsal occipito-temporal cortex in congenitally blind adults, focusing on a region of interest encompassing the hMT+ complex, which normally represents the direction of visual motion. To determine whether this region retains the same function when deprived of its usual (visual) input, an fMRI scan was taken while congenitally blind subjects heard leftward and rightward broadband noise signals, as well as static control stimuli. Wolbers et al. found that the region of interest was specifically involved in motion detection in congenitally blind subjects even though the sensory input in this case was auditory,

---

[12] Blindfolded sighted control subjects doing these auditory and tactile tasks did not show the same activation of the MOG. However, MOG activation in sighted subjects did occur in a comparable visual task.

not visual.[13] Once again, we have an impressive instance of plasticity (the fact that the dorsal visual pathway is coopted for auditory processing), but the plasticity is constrained, preserving the normal functional specificity of a dorsal pathway subregion.

Constrained plasticity has been found in the ventral visual pathway as well. For example, Striem-Amit et al. (2012) trained congenitally blind subjects and sighted controls to use a sensory substitution device, which transforms visual information from a head-mounted camera into auditory information so that soundscapes can be used to detect visual stimuli. Blind and (blindfolded) sighted subjects had up to an hour and a half of training on the device before being tested on simple geometrical shapes in different locations. Upon hearing the prompt "shape", they had to judge whether the stimulus was circular or angular; upon hearing "location", they had to judge whether it was on the left or the right side. fMRI data showed that both the blind and sighted groups exhibited differential activation for location versus shape, with activation for shape in the inferior temporal cortex (in the ventral visual stream), and activation for location in the precuneus and middle temporal sulcus/gyrus (in the dorsal visual stream). What's more, in blind subjects, there was increased activation for shape information in the ventral visual stream (in ventral Brodmann area 19), consistent with the view that prior use of the ventral visual stream for auditory stimuli enhanced this function. Thus it would appear that the large-scale functional architecture of the visual cortex—the division of labor between the dorsal and ventral streams—develops in much the same way, and with the same functions being performed in various subregions of these streams, with or without visual experience.

These and related studies showing preserved functional specificity in the visual cortex support what is known as the *metamodal hypothesis* regarding the brain's functional organization (Pascual-Leone and Hamilton, 2001). According to this hypothesis, the brain is composed of distinct computational systems whose functions are established independently of their sensory input. These systems are capable of processing sensory information from differing modalities but settle on a given modality when the input it provides is the best fit for the computations carried out—thus giving the appearance of modal-specificity. On this view, it is a misnomer to speak of the "visual cortex", the "auditory cortex", etc. Rather, each of these broad areas are composed of neural systems that engage in computations that create a preference for a given modality, but the computations performed aren't inherently about visual or auditory content and so, when the preferred input is unavailable, the brain switches to the next best fit. As our discussion above suggests, there is now a considerable amount of evidence in support of the metamodal hypothesis. For our purposes, though, what matters is the implication for nativism. Notice that to the extent that the brain is organized in this way, we have grounds to suppose that the functional specificity associated with particular regions of the brain is innate. The reason why the hMT+ computes direction of motion, for example, can't be because this is required by its visual input; it performs the same function in the complete absence of visual input in the congenitally blind. Rather, the most plausible explanation of its functional specificity is that this brain region is

---

[13] Currently there are far more studies directed to visual impairments than to other sensory impairments, but there is evidence that the same overall pattern holds when the impairment is to cortical areas that usually draw on nonvisual input. For example, work with deaf cats has found that just as the visual cortex can process auditory or tactile stimuli in the congenitally blind, the auditory cortex can process visual stimuli in the congenitally deaf. Lomber et al. (2010) examined congenitally deaf cats, which possess enhanced visual location abilities. By temporarily deactivating differing portions of auditory cortex (using surgically implanted cooling loops), Lomber et al. were able to determine that the cortical area that is responsible for this enhancement is the posterior auditory field, a region involved in the localization of *acoustic* stimuli in hearing cats.

innately organized for computing direction of motion, and this results in it selecting visual input when visual input is available because visual input is optimal for the computations it performs.

Some of the examples of innate neural systems we have been reviewing may not seem especially conceptual. However, whether they involve conceptual representations or not, they do at least contribute to the formation and processing of representations that undoubtedly are—e.g., concepts of movement, location, and spatial relations—and thereby count as part of the arrangement of special-purpose systems that explains concept acquisition for nativists. In any case, other work speaks more directly to the neural basis of conceptual-level representations.

For example, Mahon et al. (2009) examined the ventral visual stream's representation of living versus nonliving kinds. It is well known that the ventral visual stream exhibits neural specialization for these differing categories, with the representation of artifacts (e.g., tools and nonmanipulable objects) in medial regions and the representation of living animate things (e.g., animals and faces) in lateral regions. A common assumption in neuropsychology is that this medial-to-lateral organization stems from the differing visual features associated with these categories, which are claimed to lead to differing types of visual experiences with exemplars from these categories (e.g., Rogers et al. 2005). One way to test this supposition is to compare sighted and congenitally blind subjects using a common task that would be expected to generate ventral visual stream activation in sighted subjects. This is exactly what Mahon et al. did, asking sighted and congenitally blind subjects to make size judgments upon hearing words for artifacts and animals. In both blind and sighted subjects, the same medial-to-lateral organization was found. Now if representations of living and nonliving kinds were organized as they are in the ventral visual stream of sighted subjects because of a response to some measure of visual similarity, it would be deeply surprising to find the same fine-grained functional differentiation along the medial-to-lateral axis among the congenitally blind. As Mahon et al. note, the data suggest instead "that the organization of the ventral stream innately anticipates the different types of computations that must be carried out over objects from different conceptual domains" (Mahon et al. 2009, p. 403).

It is certainly noteworthy that the organization of high-level representations in the visual cortex is retained notwithstanding the complete lack of visual input. But just as important are instances of higher cognitive *amodal* neural systems retaining their functional specificity despite a lack of visual input. After all, these systems still depend on sensory information, so the information they draw upon will differ in dramatic ways when visual information is not available. Consider, for example, the impact of blindness on the development of ordinary mentalizing abilities, which include the ability to attribute mental states to others and to oneself, and the ability to reason about mental states and their role in behavior. Blind individuals lack access to many of the perceptual cues that are typically associated with these abilities, such as others' facial expressions, direction of gaze, and body posture. Blind individuals also can't rely on first-person experience to understand other people's visual perception of events. Despite these radical differences, however, the location of the neural substrates for mentalizing in early-blind individuals (including congenitally blind individuals) is the same as for sighted individuals (Bedny et al 2009). Notice how unexpected this is on the assumption that mentalizing is acquired largely on the basis of general-purpose processes that are especially sensitive to perceptual cues. Why would the same cortical areas end up with the same peculiar functions given such grossly different access to the evidence for mental activity? In contrast, this constancy in function is naturally explained on the hypothesis that these cortical areas and the functions they realize are

determined independently of perceptual input, reflecting a psychological capacity that is, to a significant extent, innate.[14]

In sum, there is considerable evidence for preserved functional specificity of neural areas even when an early sensory deficit promotes crossmodal plasticity. If the cortex were an equipotential network, the brain ought to undergo an immense functional reorganization in cases of congenital and early blindness. But that's not what happens at all. We've seen the preservation of the fine-grained functional structure in what is usually thought of as a visual-motor area, in areas that represent spatial properties (such as location and direction of motion), in the large-scale functional differentiation between the ventral and dorsal visual streams, in the functional specificity within the ventral stream (which includes different areas for representing artifacts and living kinds), and in amodal neural centers (e.g., areas associated with mentalizing) that rely heavily on sensory input. Taken together these and related studies constitute a diverse and compelling body of evidence for constrained plasticity.[15]

4.4 Constrained Plasticity in Cognitive and Conceptual Impairments

We turn now to a second type of case that argues in favor of the innate functional specificity of particular brain regions: cognitive and conceptual impairments due to focal brain damage and genetic anomalies. Such impairments argue for constrained plasticity because they show that the brain is unable to compensate for certain difficulties despite ample opportunity for neural reorganization and exposure to relevant features of the environment.

Neurological disorders typically don't affect a single functional system in isolation but rather involve a variety of co-occurring deficits. For example, a stroke may result in damage to functionally distinct yet physiologically neighboring brain areas that are equally dependent on the impeded blood flow. Nonetheless, cognitive deficits are sometimes quite specific. For example, prosopagnosia, a deficit in the ability to recognizes faces, may be accompanied by

---

[14] Bedny et al. (2009) offer a qualification to these conclusions, citing work that suggests that an understanding of false belief develops at a later age in blind individuals than sighted individuals, perhaps as late as 8 years old (e.g., Peterson et al. 2000). However, this work is based on what has come to be called *elicited-response* false-belief tasks (where children are asked direct questions regarding the mental states of others), rather than *spontaneous-response* false-belief tasks (where children display their knowledge of others' mental states in their spontaneous responses to false beliefs in others) (Baillargeon et al. 2010). Since sighted children have been shown to pass spontaneous-response tasks at much younger ages than elicited-response tasks (Onishi and Baillargeon 2005, Surian et al. 2007, Luo 2011, Kovács et al. 2010), it would be very interesting and revealing to determine if blind infants can pass non-visual, non-verbal spontaneous-response false-belief task at a comparable age to sighted infants. As far as we know, though, all spontaneous-response false-belief tasks that have been run on infants to date have been visually-based tasks, so new tasks would need to be designed in order to test this possibility.

[15] Another case of preserved functional specificity in the face of sensory deficits can be found in an area of the visual cortex that is involved in word and letter recognition in reading. The neural region subserving this ability is in fact remarkably stable across individuals and languages. Meta-analyses of numerous studies show that "the same region of the left lateral occipitotemporal sulcus always is activated, to within a few millimeters, whenever literate humans read" (Dehaene & Cohen 2011). This case is particularly interesting in light of the fact that reading is such a recent cultural invention; there couldn't be a biological adaptation for reading per se. Rather, the most plausible explanation of this functional specificity—a broadly nativist explanation—is that the abstract categorization of words and letters is always subserved by this neural region because it has an innate computational structure and functional connectivity that makes it uniquely well-suited to playing this role, and that the cultural practice of reading has been altered in history to better accommodate the peculiarities of the computational processes that it did evolve for (Dehaene 2009).

other forms of agnosia, but can also occur as a selective deficit in which the impairment is peculiar to faces. In this case, individuals may be unable to recognize altered versions of faces yet have no difficulty in recognizing comparably altered complex objects that aren't faces (Busigny et al. 2010a, 2010b). This sort of specificity regarding a representational deficit can persist in spite of many years of exposure to relevant stimuli and a strong vested interest in the subject domain. Individuals who suffer from prosopagnosia, for example, will often have as much difficulty recognizing faces of familiar and emotionally significant people in their lives (parents, children, partners) as they have with strangers. In addition, *category-specific deficits* in semantic memory (memory related to general knowledge) aren't tied to any particular type of task or a given modality (Capitani et al. 2003). A selective deficit for the category of living kinds, for instance, may show up equally in an inability to recognize animals in a picture-naming task and in purely verbal queries about the features of different animals.

The specificity of a category-specific deficit might be explained in a number of different ways. We will consider two broad classes of explanation regarding deficits in semantic memory, focusing on the well-studied example of category-specific deficits in the representation of living kinds—e.g., patients with significant impairments for animals (elephant, duck, etc.) in contrast with artifacts (pen, key, etc.). The standard empiricist explanation of such cases holds that semantic memory isn't organized in terms of a categorical distinction between the living and the nonliving (or animal vs. artifact), but instead is organized in terms of the properties that exemplars of particular categories possess. Different types of properties are taken to figure more prominently in the representation of categories of living versus nonliving kinds. For instance, on one influential account, visual properties are taken to be more prominent for living kinds, and functional properties for nonliving kinds. If this account were correct, then focal damage to the neural substrate for the representation of visual properties would disproportionately affect living kinds, while damage to the representation of functional properties would disproportionately affect nonliving kinds (Warringon and McCarthy 1983, Farah and McClelland 1991). The nativist approach, in contrast, maintains that semantic memory *is* organized in terms of a categorical distinction between living and nonliving kinds, and that in general there are innately dedicated neural circuits related to a number of fundamental category types with particular evolutionary significance, such as animals, tools, faces, and food (Caramazza and Shelton 1998, Mahon and Caramazza 2009).

As our interest is in the empiricism-nativism debate, a particularly important type of case to consider in evaluating these two different types of explanation is one in which a category-specific deficit results from neural damage or from a genetic disorder that affects early development. Farah and Rabinowitz (2003) documented the case of Adam, who sustained brain damage when he was just one day old. At age 16, Adam was tested for his knowledge of living and nonliving kinds, and a significant difference between the two was found. Adam had a severe impairment for knowledge regarding living kinds (responding to testing at chance levels), yet his performance was normal or near normal regarding nonliving kinds. His difficulty with living kinds was also comprehensive in that it affected visual and non-visual properties alike, while his knowledge of nonliving kinds (both visual and non-visual) was spared. Consequently, Adam's psychological profile is inconsistent with the empiricist explanation of category-specific deficits in terms of selective damage to the representation of a given type of property (in this case, to visual properties).

What's more, Adam's case speaks directly to the limitations on neural plasticity in cognitive development. Despite the fact that the neural damage occurred very early in development, and

despite the fact that Adam had years of experience in infancy and childhood in which other aspects of his psychological development proceeded normally, his brain wasn't able to compensate for the damage it had sustained. As Farah and Rabinowitz put it:

> … phrased in terms of Adam's surviving brain tissue, despite its adequacy for acquiring semantic memory about nonliving things, it could not take over the function of semantic memory for living things. This implies that prior to any experience with living and nonliving things, we are destined to represent our knowledge of living and nonliving things with distinct neural substrates. This in turn implies that the distinction between living and nonliving things, and the anatomical localization of knowledge of living things, are specified in the human genome. (Farah and Rabinowitz 2003, p. 408)

In a related study, Farah et al. (2000) examined a different specific representational deficit in the same subject, namely, Adam's difficulty with faces. At the age of sixteen, Adam had the classic profile of prosopagnosia—lesions in occipitotemporal cortex (bilaterally), with a severe impairment in the ability to recognize faces relative to good, though not perfect, object recognition abilities.[16] As with the living/nonliving distinction, this uneven cognitive profile raises the questions of why other neural tissue was unable to compensate for the damaged neural tissue—a striking lack of plasticity—especially given the obvious importance of face recognition in daily life.

Now there are a number of possible explanations for why the representation of faces might be impaired, just as there are different possible explanations for the selective impairment to the representation of living (or nonliving) kinds. Duchaine et al. (2006) addressed this issue by examining another patient, Edward, who suffered from developmental prosopagnosia. In this study, Duchaine et al. took advantage of the opportunity to test on a single subject all of the alternatives to the nativist domain-specific explanation that have appeared in the face-perception literature. Among the empiricist explanations Duchaine et al. looked into were the possibility that Edward suffered from a general difficulty regarding the representation of individuals within a category, a general difficulty with holistic processing, a general difficulty with configural processing (i.e., representing the spacing between features), and a general difficulty in acquiring expertise for object categories. For example, the configural processing explanation was evaluated by having Edward make same-different judgments for photographs of faces and houses that had been digitally altered. The distance between the eyes or windows was changed, or these features themselves were replaced with similar features in the same relative spacing. In this case, Edward's performance was normal for detecting changes to houses, but three standard deviations below the mean for detecting commensurate changes to faces. Likewise, the expertise hypothesis was evaluated using corresponding face- and body-matching tests, in which the goal was to identify which of two rotated faces or headless bodies matched a target. Here too Edward had great difficulty with faces, but his performance with bodies was normal—in fact, he scored in the high end of the normal range for body recognition. These and the results from Duchaine et al.'s other tests indicate that Edward's difficulty is genuinely face-specific, and consequently that there are face-specific developmental mechanisms that may be selectively impaired.

---

[16] "In everyday life he is unable to recognise faces, whereas his object recognition ability is fully adequate for activities of daily living" (Farah et al. 2000, p. 122).

Edward's impairment (unlike Adam's) is most likely the result of a genetic anomaly.[17] Though not all genetic disorders that result in representational deficits are as focused as prosopagnosia—most result in uneven but predictable profiles of spared conceptual abilities and impairments—they can still provide an excellent source of evidence regarding the limits on the brain's plasticity. For example, individuals with Williams' Syndrome, a rare genetic disorder (Schubert 2009), show severe deficits in certain types of reorientation tasks which rely on geometrical representation[18] but relatively spared face recognition abilities (Lakusta et al. 2010, Bellugi et al. 2000), and have intact biological motion representation in spite of other types of motion representation deficits (Jordan et al. 2002, Reiss et al. 2005).

One particularly well-studied and illuminating case is the impairment to mentalizing abilities found in individuals with Autism Spectrum Disorder (ASD).[19] In a classic early investigation, Baron-Cohen et al. (1985) examined three groups of children on a false-belief task—clinically normal preschool children, children with Down's syndrome, and high-functioning children with ASD. In Baron-Cohen et al.'s test for the representation of false belief (the Sally-Anne task), subjects witness a protagonist (Sally) place a marble in her basket only to have it moved (by Anne) to a box while she is away from the scene. When Sally returns, subjects are asked where she will look for her marble. Baron-Cohen et al. found that clinically normal children and Down's Syndrome children both answered correctly, saying that she will look in the basket (where Sally should falsely think that it is), while children with ASD overwhelmingly gave the incorrect response, saying that she will look in the box (where it actually is). Subsequent work by Leslie and Thaiss (1992) showed that this failure is specific to the understanding of belief and not part of a general difficulty with understanding representation. Leslie and Thaiss compared clinically normal preschool children with high-functioning children with ASD—this time using both false belief tasks and structurally similar tasks with photographs and maps. (In a false photograph task, for example, a Polaroid photo is taken of an object in one location, only to have the object moved before the photo is developed. Then the question asked is where the object will be in the photograph.) Leslie and Thaiss found that children with ASD who failed false-belief tasks were able to pass false-photograph or false-map tasks. By contrast, children who didn't have ASD found the false-photograph and false-map tasks more difficult than the false-belief task.

More recent work has found that children with ASD not only have difficulties with elicited-response false-belief tasks (in which they are explicitly asked to respond to a false belief scenario), but they are also unable to anticipate an actor's actions when presented with evidence of the actor's false belief in a spontaneous-response task (Senju et al. 2010). This is not due to a general inability to understand action, as they correctly predict an agent's actions when the agent doesn't have a false belief, and are able to correctly attribute goals to an agent even when the

---

[17] Face recognition ability is highly heritable. Polk et al. (2007) compared the patterns of neural activity to faces and other stimuli in monozygotic and dizygotic siblings and found that the activity was significantly more similar for monozygotic siblings for faces and places, but not for pseudowords or chairs. Likewise, Wilmer et al. (2010) tested monozygotic and dizygotic twins on a face memory task and several control tasks for nonface memory. They found that "genetic differences can account for most of the stable variation in face recognition ability in healthy adults" and that this is specific to faces (Wilmer et al. 2010, p. 5239).

[18] In a typical reorientation task, subjects in a rectangular room are shown the hiding place for an object and are gently spun around until they become disoriented. They are then asked to locate the object, which requires using the geometry of the room to become reoriented.

[19] Hereditability studies indicate a strong genetic component in ASD, but it appears that there are numerous different genetic anamolies that give rise to the characteristic impairments in ASD (see, e.g., Huguet et al. 2013).

agent fails to achieve his goal (Carpenter et al. 2001). Further, this sort of impairment persists into adulthood. Senju et al. (2009) found that adults with ASD who can correctly answer explicit questions about what an agent with false beliefs will do nonetheless fail to spontaneously anticipate that an agent will act the same way in a live situation. This suggests that they are solving elicited-response tasks using consciously formulated rules that substitute for an intuitive understanding of the source of action. Other work suggests a similar conclusion. Ordinarily people modulate their behavior when they are observed because of the potential effect on their social reputation (e.g., giving more money to a charity in the presence of others than when alone). In contrast, high-functioning adults with ASD don't modulate their behavior in this way (Izuma et al. 2011). Likewise, ordinarily people take into account the absence of negative intentions when formulating a moral judgment pertaining to someone who accidentally causes a negative outcome. Here too high-functioning adults with ASD behave differently, treating cases with and without negative intentions equally (Moran et al 2011).[20] Thus, there is a convergence of evidence that suggests that ASD is associated with a selective representational impairment, one that affects the formation and use of certain mental state concepts but not other concepts of comparable difficulty.

4.5 Summary and the Future

We began this section by reexamining the argument against nativist views of concepts that appeals to considerations having to do with neural plasticity and showed that this argument isn't so compelling after all. We then presented a range of different types of evidence for the opposing nativist hypothesis of constrained plasticity:

1. *Evidence from preserved neural structural organization in the absence of relevant environmental input.* This was seen in the genetically altered mice, whose brains developed the same structural organization as their normal littermates even though the genetic manipulations disrupted all synaptic transmission, eliminating any possibility for sensory information or feedback to affect development.

2. *Evidence from preserved neural functional specificity despite early sensory deprivation.* This was seen, for example, in the preservation of the large-scale functional organization of the visual system in the congenitally blind (the division of labor inherent to the ventral and dorsal visual streams), and in cases where the functional/representational specificity of particular neural subregions remains the same even though the input is no longer visual (e.g., the assignment of the same neural region to the abstract representation of living kinds).

3. *Evidence from conceptual and cognitive deficits resulting from early focal neurological trauma.* In this case, we saw that the brain's plasticity is unable to compensate for such deficits despite ample opportunity for neural reorganization and exposure to relevant

---

[20] The right temporo-parietal junction (rTJP), which is known to be a critical mentalizing brain area (Koster-Hale and Saxe 2013), is particularly involved in modulating moral judgments according to whether a harm is accidental or intentional (Buckhotz et al. 2008, Young and Saxe 2009). Interestingly, the normal spatially distinct responses within the rTJP for accidental vs. intentional harms is absent in adults with ASD (Koster-Hale et al. 2013).

features of the environment (e.g., Adam's inability to represent faces or living kinds in adulthood).

4. *Evidence from conceptual and cognitive deficits resulting from genetic anomalies*. The brain's plasticity is also unable to compensate for deficits resulting from developmental genetic anomalies (e.g., spatial deficits associated with Williams Syndrome and mentalizing deficits associated with autism) despite ample opportunity for neural reorganization and exposure to relevant features of the environment.

We conclude that the neural plasticity of the brain takes the form of *constrained* plasticity—development that is not open-endedly plastic, but instead is highly constrained in ways that suggest important innate biases, predispositions, and limits on structure and function.[21]

We want to reiterate, however, that the case for concept nativism doesn't stand entirely with the argument from neural wiring. While what has been discovered about the details of neural structure and function strongly supports the nativist viewpoint, equally important are the findings in such fields as animal psychology, developmental psychology, evolutionary psychology, linguistics, and cross-cultural studies. Moreover, the connection between work in these (and other) areas of cognitive science and the argument from neural wiring is a relatively unexplored and potentially rich source of insight about the origins of human concepts. For this reason, we'd suggest that an exciting and important direction for future work on concepts is to revisit the argument from neural wiring with the aim of looking for a far more extensive body of innate systems of representation, guided by nativist hypotheses that are independently motivated in these other areas of cognitive science.

5. Conclusion

Concept nativism holds that a significant number of concepts are either innate or acquired via innate special-purpose acquisition systems. Concept nativism isn't opposed to learning; rather, it offers a distinctive perspective on learning, one that is grounded in the idea that much learning takes place only because the mind is innately structured to extract specific types of information from the world and to process this information in particular ways. Concept nativism tells us that we shouldn't assume that there is a single general-purpose acquisition mechanism that accounts for the origin of our concepts for animals, artifacts, mental states, individual people, and so on. These and other content domains may well depend on distinct special-purpose innate acquisition systems that are geared towards specific types of content.

The case for concept nativism takes the form of an inference to the best explanation that draws on evidence from a range of disciplines. What we have shown here is that one of the major objections to nativism—its alleged neurological implausibility in the face of neural plasticity—is unfounded. Neural plasticity isn't as flexible and open-ended as nativism's critics have supposed. On the contrary, the plasticity of the brain is highly constrained in a way that argues *for* concept nativism. The brain is not comprised of an equipotential network that is sculpted into

---

[21] This list of different types of evidence for the hypothesis of constrained plasticity is not intended to be exhaustive. Though space considerations prohibit an exploration of additional types of evidence, two that are worth mentioning are (1) evidence from twin studies (see, e.g., footnote 17 above) and (2) evidence for parallel neurological structures subserving the same functions across species (see, e.g., Kriegeskorte et al. 2008).

differentiated functional units through sensory experience. Rather, the brain is innately differentiated into a complex arrangement of distinct neural systems specialized for processing specific types of information, exactly as concept nativism predicts.

References

Arno, P., De Volder, A.G., Vanlierde, A., Wanet-Defalque, M.C., Streel, E., Robert, A., Sanabria-Bohorquez, S., Veraart, C. (2001). Occipital activation by pattern recognition in the early blind using auditory substitution for vision. *Neuroimage* 13, 632--645.

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in cognitive sciences*, 14(3), 110–118.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind?" *Cognition*, 21(1), 37–46.

Bedny, M., Pascual-Leone, A., & Saxe, R. R. (2009). Growing up blind does not change the neural bases of Theory of Mind. *Proceedings of the National Academy of Sciences of the United States of America*, 106(27), 11312–11317.

Bellugi, U., Lichtenberger, L., Jones, W., Lai, Z., & St George, M. (2000). I. The neurocognitive profile of Williams syndrome: a complex pattern of strengths and weaknesses. *Journal of Cognitive Neuroscience*, 12(Supplement 1), 7–29.

Boyd, R. and Richerson, P. (1985). *Culture and the Evolutionary Process*. The University of Chicago Press.

Bregman, A.S., & Pinker, S. (1978). Auditory streaming and the building of timbre. *Canadian Journal of Psychology*, 32(1), 19-31.

Buckholtz, J. W., Asplund, C. L., Dux, P. E., Zald, D. H., Gore, J. C., Jones, O. D., & Marois, R. (2008). The Neural Correlates of Third-Party Punishment. *Neuron*, 60(5), 930–940.

Buller, D.J., & Hardcastle, V. (2000). Evolutionary psychology, meet developmental neurobiology: Against promiscuous modularity. *Brain and Mind*, 1(3), 307–325.

Busigny, T., Graf, M., Mayer, E. N., & Rossion, B. (2010a). Acquired prosopagnosia as a face-specific disorder: Ruling out the general visual similarity account. *Neuropsychologia*, 48(7), 2051–2067.

Busigny, T., Joubert, S., Felician, O., Ceccaldi, M., & Rossion, B. (2010b). Holistic perception of the individual face is specific and necessary: Evidence from an extensive case study of acquired prosopagnosia. *Neuropsychologia*, 48(14), 4057–4092.

Capitani, E., Laiacona, M., Mahon, B.Z., and Caramazza, A. (2003). What are the facts of category-specific deficits? A critical review of the clinical evidence. *Cognitive Neuropsychology*, 20, 213–261.

Caramazza, A., and Shelton, J.R. (1998). Domain-specific knowledge systems in the brain the animate-inanimate distinction. *Journal of Cognitive Neuroscience*, 10, 1–34.

Carpenter, M., Pennington, B.F., & Rogers, S.J. (2001). Understanding of others' intentions in children with autism. *Journal of Autism and Developmental Disorders*, 31 (6), 589-599.

Chomsky, N. (1991). Linguistics and cognitive science: problems and mysteries. In A. Kasher (ed.), *The Chomskyan Turn*. Blackwell.

Cohen, L.G., Celnik, P., Pascual-Leone, A., Corwell, B., Faiz, L., Dambrosia, J., Honda, M., Sadato, N., Gerloff, C., Catalá, M.D., & Hallett, M. (1997). Functional relevance of cross-modal plasticity in blind humans. *Nature*, 389(6647), 180–183.

Collignon, O., Lassonde, M., Lepore, F., Bastien, D., and Veraart, C. (2007). Functional cerebral reorganization for auditory spatial processing and auditory substitution of vision in early blind subjects. *Cerebral Cortex* , 17, 457–65.

Csibra, G. & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society, B*, 366, 1149-1157.

Curtiss, S., & Schaeffer, J. (2005). Syntactic development in children with hemispherectomy: The I-, D-, and C-systems. *Brain and Language*, 94(2), 147–166.

Curtiss, S., & de Bode, S. (2003). How normal is grammatical development in the right hemisphere following hemispherectomy? The root infinitive stage and beyond. *Brain and Language*, 86(2), 193–206.

Danelli, L., Cossu, G., Berlingeri, M., Bottini, G., Sberna, M., & Paulesu, E. (2013). Is a lone right hemisphere enough? Neurolinguistic architecture in a case with a very early left hemispherectomy. *Neurocase*, 19(3), 209–231.

Dehaene, S. (2009) *Reading in the Brain*. Viking.

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in cognitive sciences*, 15(6), 254–262.

Duchaine, B. C., Yovel, G., Butterworth, E. J., & Nakayama, K. (2006). Prosopagnosia as an impairment to face-specific mechanisms: Elimination of the alternative hypotheses in a developmental case. *Cognitive Neuropsychology*, 23(5), 714–747.

Elman, J.L., Bates, E.A., Johnson, M.H., Karmiloff-Smith, A. Parisi, D., Plunkett, K. (1996) *Rethinking Innateness: A Connectionist Perspective on Development*. MIT Press.

Farah, M. J., & McClelland, J. L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General,* 120(4), 339–357.

Farah, M. J., & Rabinowitz, C. (2003). Genetic and environmental influences on the organisation of semantic memory in the brain: Is "living things" an innate category? *Cognitive Neuropsychology*, 20(3-6), 401–408.

Farah, M. J., Rabinowitz, C., Quinn, G. E., & Liu, G. T. (2000). Early commitment of neural substrates for face recognition. *Cognitive Neuropsychology*, 17(1-3), 117–123.

Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences*, 108(39), 16428–16433.

Fodor, J.A. (1975). *The Language of Thought*. Tomas Y. Crowell.

Fodor, J.A. (1981). The present status of the innateness controversy. In his *Representations: Philosophical Essays on the Foundations of Cognitive Science*. MIT Press.

Fodor, J. (1998). *Concepts: Where Cognitive Science Went Wrong*. Oxford University Press.

Fodor, J.A. (2008). *LOT2: The Language of Thought Revisited*. Oxford University Press.

Gazzaniga, M.S., Irvy, R.B., & Mangun, G.R. (2009). *Cognitive Neuroscience: The Biology of the Mind*, 3rd ed. W. W. Norton & Co.

He, S., Liu, H., Jiang, Y., Chen, C., Gong, Q., & Weng, X. (2009). Transforming a left lateral fusiform region into VWFA through training in illiterate adults. *Journal of Vision*, 9(8), 853–853.

Huguet, G., Ey, E., & Bourgeron, T. (2013). The Genetic Landscapes of Autism Spectrum Disorders. *Annual Review of Genomics and Human Genetics*, 14(1), 191–213.

Izuma, K., Matsumoto, K., Camerer, C. F., & Adolphs, R. (2011). Insensitivity to social reputation in autism. *Proceedings of the National Academy of Sciences of the United States of America*, 108(42), 17302–17307.

Jordan, H., Reiss, J. E., Hoffman, J. E., & Landau, B. (2002). Intact perception of biological motion in the face of profound spatial deficits: Williams syndrome. *Psychological Science*, 13(2), 162–167.

Koster-Hale, J., Saxe, R. (2013) Functional Neuroimaging of Theory of Mind. In S. Baron-Cohen, M. Lombardo & H. Tager-Flusberg (eds.) *Understanding Other Minds*, 3rd Edition. Oxford University Press.

Koster-Hale, J., Saxe, R., Dungan, J., & Young, L.L. (2013). Decoding moral judgments from neural representations of intentions. *Proceedings of the National Academy of Sciences*, 110(14), 5648–5653.

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, 330(6012), 1830–1834.

Kriegeskorte, N., Mur, M., Ruff, D.A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., & Bandettini, P.A. (2008). Matching Categorical Object Representations in Inferior Temporal Cortex of Man and Monkey. *Neuron*, 60(6), 1126–1141.

Lakusta, L., Dessalegn, B., & Landau, B. (2010). Impaired geometric reorientation caused by genetic defect. *Proceedings of the National Academy of Sciences*, 107(7), 2813–2817.

Laurence, S., & Margolis, E. (2002). Radical concept nativism. *Cognition*, 86, 25-55.

Laurence, S., & Margolis, E. (in prep). *The Building Blocks of Thought*.

Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition*, 43(3), 225–251.

Lingnau, A., Strnad, L., He, C., Fabbri, S., Han, Z., Bi, Y., & Caramazza, A. (2012 advance access). Cross-modal plasticity preserves functional specialization in posterior parietal cortex. *Cerebral Cortex*.

Lomber, S. G., Meredith, M. A., & Kral, A. (2010). Cross-modal plasticity in specific auditory cortices underlies visual compensations in the deaf. *Nature Neuroscience*, 13(11), 1421–1427.

Luo, Y. (2011). Do 10-month-old infants understand others' false beliefs? *Cognition*, 121(3), 289–298.

Mahon, B. Z., Anzellotti, S., Schwarzbach, J., Zampini, M., & Caramazza, A. (2009). Category-specific organization in the human brain does not require visual experience. *Neuron*, 63(3), 397–405.

Mahon, B. Z., & Caramazza, A. (2009). Concepts and categories: A cognitive neuropsychological perspective. *Annual Review of Psychology*, 60(1), 27–51.

Majewska, A. K., & Sur, M. (2006). Plasticity and specificity of cortical processing networks. *Trends in Neurosciences*, 29(6), 323–329.

Marcus, G. (2004). *The Birth of the Mind: How a Tiny Number of Genes Creates the Complexities of Human Thought*. Basic Books.

Moran, J. M., Young, L. L., Saxe, R., Lee, S. M., O'Young, D., Mavros, P. L., & Gabrieli, J. D. (2011). Impaired theory of mind for moral judgment in high-functioning autism. *Proceedings of the National Academy of Sciences of the United States of America*, 108(7), 2688–2692.

Onishi, K.H. & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science* 308, 255-258.

Pascual-Leone, A., & Hamilton, R. (2001). The metamodal organization of the brain. *Progress in Brain Research*, 134, 427–445.

Peterson, C. C., Peterson, J. L., & Webb, J. (2000). Factors influencing the development of a theory of mind in blind children. *British Journal of Developmental Psychology*, 18(3), 431–447.

Pinker, S. (2002). *The Blank Slate: The Modern Denial of Human Nature*. Penguin.

Pinker, S. (2007). *The Stuff of Thought*. Viking.

Polk, T.A., Park, J., Smith, M.R., and Park, D.C. (2007). Nature versus nurture in ventral visual cortex: a functional magnetic resonance imaging study of twins. *The Journal of Neuroscience*, 27, 13921–13925.

Prinz, J. (2005). The return of concept empiricism. In H. Cohen & C. Leferbvre (eds.) *Handbook of Categorization in Cognitive Science* (pp. 679-95). Elsevier Science Ltd.

Prinz, J. (2012). *Beyond Human Nature: How Culture and Experience Shapes Our Lives*. Allen Lane.

Proulx, M. J., Brown, D. J., Pasqualotto, A., & Meijer, P. (in press). Multisensory perceptual learning and sensory substitution. Neuroscience & Biobehavioral Reviews, 1–10.

Quartz, S.R. (2003). Innateness and the brain. *Biology & Philosophy*, 18(1), 13–40.

Quine, W.V.O. (1969). Linguistics and philosophy. In S. Hook (*ed.*) *Language and Philosophy: A Symposium*, New York: NYU Press.

Reich, L., Szwed, M., Cohen, L., & Amedi, A. (2011). A Ventral Visual Stream Reading Center Independent of Visual Experience. *Current Biology*, 21(5), 363–368.

Reiss, J. E., Hoffman, J. E., & Landau, B. (2005). Motion processing specialization in Williams syndrome. *Vision Research*, 45(27), 3379–3390.

Renier, L.A., Anurova, I., De Volder, A.G., Carlson, S., VanMeter, J., & Rauschecker, J.P. (2010). Preserved Functional Specialization for Spatial Processing in the Middle Occipital Gyrus of the Early Blind. *Neuron*, 68(1), 138–148.

Rogers, T.T. & McClelland, J.L (2004). *Semantic Cognition: A Parallel Distributed Processing Approach*. MIT Press.

Rogers, T. T., Hocking, J., Mechelli, A., Patterson, K., & Price, C. (2005). Fusiform activation to animals is driven by the process, not the stimulus. *Journal of Cognitive Neuroscience*, 17(3), 434–445.

Sadato, N., Pascual-Leone, A., Grafman, J., Ibañez, V., Deiber, M.P., Dold, G., & Hallett, M. (1996). Activation of the Primary Visual Cortex by Braille Reading in Blind Subjects, *Nature,* 380, 526–528.

Schubert, C. (2009). The genomic basis of the Williams–Beuren syndrome. *Cellular and Molecular Life Sciences*, 66(7), 1178–1197.

Senju, A., Southgate, V., Miura, Y., Matsui, T., Hasegawa, T., Tojo, Y., Osanai, H., & Csibra, G. (2010). Absence of spontaneous action anticipation by false belief attribution in children with autism spectrum disorder. *Development and Psychopathology*, 22(02), 353.

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind eyes: An absence of spontaneous theory of mind in Asperger Syndrome. *Science*, 325(5942), 883–885.

Sharma, J., Angelucci, A., & Sur, M. (2000). Induction of visual orientation modules in auditory cortex. *Nature*, 404(6780), 841–847.

Sripada, C. S., & Stich, S. (2006). A framework for the psychology of norms. In P. Caruthers, S. Laurence & S. Stich (eds.) *The Innate Mind: Culture and Cognition*, pp. 280-301. Oxford University Press.

Stich, S. (1975). Introduction: The idea of innateness. In S. Stich (ed.) *Innate Ideas*. University of California Press.

Striem-Amit, E., Cohen, L., Dehaene, S., & Amedi, A. (2012). Reading with Sounds: Sensory Substitution Selectively Activates the Visual Word Form Area in the Blind. *Neuron*, 76(3), 640–652.

Striem-Amit, E., Dakwar, O., Reich, L., & Amedi, A. (2012). The large-Scale Organization of "Visual" Streams Emerges Without Visual Experience. *Cerebral Cortex*, 22(7), 1698–1709.

Sur, M. (1988). Visual plasticity in the auditory pathway: visual inputs induced into auditory thalamus and cortex illustrate principles of adaptive organization in sensory systems. In M.A. Arbib & S.-i. Amari (eds.) *Dynamic Interactions in Neural Networks: Models and Data,* (pp. 35-51). Springer-Verlag.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, 18(7), 580–586.

Verhage, M., Maia, A.S., Plomp, J.J., Brussaard, A.B., Heeroma, J.H., Vermeer, H., Toonen, R.F., Hammer, R.E., van den Berg, T.K., Missler, M., Geuze, H.J., & Südhof, T.C. (2000). Synaptic assembly of the brain in the absence of neurotransmitter secretion. *Science*, 287(5454), 864–869.

von Melchner, L., Pallas, S.L. & Sur, M. (2000). Visual behaviour mediated by retinal projections directed to the auditory pathway. *Nature*, 404, 871-876.

Warrington, E.K., McCarthy, R.A. (1983). Category specific access dysphasia. *Brain*, 106, 859–878.

Wilmer, J.B., Germine, L., Chabris, C.F., Chatterjee, G., Williams, M., Loken, E., Nakayama, K., & Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of Sciences*, 107(11), 5238–5241.

Wolbers, T., Zahorik, P., & Giudice, N. A. (2011). Decoding the direction of auditory motion in blind humans. *NeuroImage*, 56(2), 681–687.

Young, L., & Saxe, R. (2009). Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia*, 47(10), 2065–2072.