„Analiza i Egzystencja" 24 (2013)
ISSN 1734-9923

JUSTYNA JAPOLA-DESVERGNES*

NO MORE MAD-DOG CONCEPT NATIVISM

Abstract

According to Jerry Fodor, there are two possibilities with respect to the origin of concepts: a concept is either innate, or it is acquired from experience by means of the process of 'concept learning.' Fodor's view in his 1975 and 1980 is that the empiricist method of concept acquisition, forming and testing hypotheses about objects that fall under a concept, can only work for complex concepts. He claims that we must possess some concepts in order to form hypotheses, so none of our simple (or primitive) concepts can be learned. If we have them—then they must be innate. In his 2008 Fodor contends furthermore that *no* new concepts at all, neither primitive, nor complex, can be learned, and so that *all* concepts must be innate. I argue that Fodor's famous mad-dog concept nativism should be rejected because it could only work together with what I call neural nativism, and the latter turns out to be scientifically untenable. In addition, I suggest that one's position in the empiricism/nativism debate should be a function of one's account of the architecture of the mind, which, in Fodor's case, only implies

* Justyna Japola-DesVergnes was born in Lublin, Poland. She has a Master's Degree in French Language and Literature from the Catholic University of Lublin, and a Ph.D. in Philosophy from Georgetown University. Dr. Japola-DesVergnes' areas of interest include analytic Thomism, philosophy of mind, philosophy of religion, Medieval philosophy, and French literature.

architectural nativism, and so, the existence of innate *mechanisms*. It neither obviously requires, nor precludes, representational—or concept—nativism.

Keywords: Fodor, concepts, concept acquisition, concept nativism, architecture of the mind

1. Introduction

Is the human cognitive endowment innate? And in particular, do we have innate concepts? According to Jerry Fodor,[1] there are two possibilities with respect to the origin of concepts: a concept is either innate or it is acquired from experience by means of the process of 'concept learning.' Since, as Fodor argues, it turns out that no new (primitive) concepts can be learned, we must conclude that all (primitive) concepts are innate.[2]

This is Fodor's famous 'mad-dog' concept nativism,[3] a view that not many philosophers find plausible, but few dare to attack (cf. Cowie 1998, p. 227). I argue that Fodor's radical nativism about concepts should be rejected based on neurological evidence. I also show that the reason why Fodor ends up a concept nativist is that he assumes that if a trait is not learned, then it must be triggered, and that it can be triggered only if it is innate. Once we reject the latter unsupported assumption, and provide a plausible definition of innateness, it turns out that Fodor himself has no reason to hold on to his radical nativism about concepts.

2. Concepts and concept acquisition according to Fodor

According to Fodor, a subject is said to have a concept C if she is able to think about something *as (a) C* (if she can bring the property C before her mind; cf. Fodor 2008, p. 138). Concepts, for Fodor, are thought-parts. They are mental entities (species of representations) that are constituents of cognitive mental states, i.e., of propositional attitudes, such as beliefs, desires, thoughts

[1] My main sources are Fodor's (1975; 1981; 1998 and 2008).

[2] As I explain below, in his 2008 Fodor concludes that no new concept at all, neither primitive, nor complex, can be learned, and so that all concepts must be innate.

[3] The term first appears in Fodor's (1984, p. 39).

and the like (Fodor 1998, p. 6, and 2000, pp. 14–15). *Lexical* concepts, such as DOG (or GREEN, or BACHELOR), are expressible by morphologically simple predicate terms. Some lexical concepts are *primitive*: they do not have an internal structure and cannot be decomposed into other concepts.[4] The set of primitive lexical concepts of any natural language is finite (cf. Fodor1981, p. 261). *Phrasal* concepts, such as the concept LIVES IN CHICAGO AND EATS MANGLEWORTS, are expressible by morphologically *complex* predicate terms.

The semantic properties of phrasal concepts depend on the semantic properties of lexical concepts (the latter are constituents of the former). Phrasal concepts are acquired by the application of the recursive constructive procedures to primitive lexical concepts (there are, thus, infinitely many phrasal concepts). Lexical concepts, according to Fodor, either are acquired as a result of what is called the process of *concept learning*, or they are *innate* and only *triggered* by experience.

Concept learning vs. triggering

Concept learning is an epistemic process: it involves the formation of beliefs and the search for evidence. Fodor explains:

> the mechanisms of concept learning are realizations of some species of inductive logic. In particular, they involve the formulation and confirmation of hypotheses about the identity of the concept being learned. (Fodor 1981, p. 267; cf. also Fodor 1975, p. 42)
>
> Each trial provides inductive evidence pro or con a hypothesis of the form: *the concept being learned is the concept of something which is…*, where what goes in the blank is a candidate specification of those properties of a stimulus in virtue of which it satisfies the concept. (Fodor 1981, p. 268)

A crucial and distinguishing characteristic of concept learning, Fodor explains, is a rational connection between the concept and what it represents (Fodor 1981, p. 275). When a concept is acquired in a *rational-causal* way, "the experiences which eventuate in the availability of such a concept are

[4] See footnote 6 below. According to some interpretations, for Fodor all lexical concepts are primitive (cf. Jackendoff 1989, p. 98).

held to bear a confirmation relation to some hypothesis which specifies the internal structure of the concept" (Fodor 1981, p. 272). In other words, in the case of those concepts which are learned, it is possible to provide an intentional description of the acquisition process and to show that the acquisition of the concept is rationally related to the experiences that gave rise to it, that it 'makes sense' in the light of those experiences.

Concepts which are *triggered* are acquired not by learning, but by "opening one's eyes and looking" (Fodor 1975, p. 97). All that is needed for the acquisition of a concept in this case is a properly functioning sensorium and the right kind of stimulus. As Fodor explains, "the structure of the sensorium is such that certain inputs trigger the availability of certain concepts" (Fodor 1981, p. 273). The connection between a concept and the experience that triggers it is *brute-causal*. The experiences of the organism *do not* stand in a confirmation relation to the concepts whose availability they occasion (Fodor 1981, p. 273), but rather they "function as the innately specified triggers" (Fodor 1981, p. 280).

Standard Argument

In what he calls the Standard Argument,[5] Fodor shows at first that it turns out to be impossible to learn any *new* primitive lexical concept by means of concept learning, and thus that all of our primitive concepts must be innate. Here is the gist of Fodor's reasoning. Imagine that you are a subject in an experiment in which your task is to learn a new concept: FLURG. The experimenter shows you different cards with colored geometrical figures on them. You make guesses whether the card you see has a flurg on it or not. You make a first hypothesis: perhaps there is a flurg on a card with a yellow square. The experimenter gives a negative answer. Your hypothesis is disconfirmed. You hypothesize that flurg may be a green circle. This gets you a positive answer from the experimenter. The hypothesis is confirmed. Eventually you arrive at the right answer: a flurg is a figure that is either green or a triangle.

It is clear that you were able to complete the task only because you already had the two concepts, GREEN and TRIANGLE, at your disposal. But if that's the case, then you did not learn any new primitive concept. The

[5] See Fodor (1998, p. 123ff; 1981, pp. 266–269).

concepts you used to figure out what a flurg is were already there in your conceptual apparatus. You were able to successfully formulate and confirm the hypothesis that a flurg is something green or triangular because you already had both concepts GREEN and TRIANGLE. And so, it turns out that what is called ‘concept learning’ does not really refer to any process by means of which we would *learn* any *new* primitive concepts. Hypotheses can only be formed using some preexisting concepts. You cannot *acquire* DOG by forming hypotheses about dog-like objects because in order to form such hypotheses you would need to already have the very concept that you are supposed to learn (cf. Fodor 1981, pp. 269, 272). And since primitive concepts cannot be learned, they must be innate.[6]

In his 1975 and 1998, Fodor states that by means of ‘concept learning’ it is possible to acquire complex (i.e., non-primitive) concepts. And so, even though a primitive concept RED cannot be learned because it is primitive, the concept FLURG (which means: GREEN OR TRIANGULAR) could be learned. FLURG cannot be identified with either of its constituents; GREEN and TRIANGULAR must be previously possessed, and FLURG can be learned as a construction from these two previously possessed basic concepts. In his 2008, however, Fodor notices that his argument works equally well for complex concepts, and so, that “the whole notion of concept learning is per se confused” (Fodor 2008, p. 130). As an illustration, Fodor considers the concept GREEN AND TRIANGULAR, and concludes that it cannot be learned by the method of hypothesis testing, because we would need to possess this very concept together with its constituents before making any hypotheses. The explanation of this fact appears already in his 1975, where Fodor says:

> The hypothesis whose acceptance is necessary and sufficient of learning [the concept] C is that C is that concept which satisfies the individuating conditions on Φ for some or other concept Φ. But, trivially, a concept that satisfies the conditions which individuate Φ *is* the concept Φ. It follows that no process which consists of confirming such a hypothesis could be

[6] Another element in Fodor’s argument for his mad-dog nativism is informational atomism, that is, a view according to which most, if not all, of our lexical concepts are primitive: they are atoms, unstructured symbols of the language of thought, carrying information about features of the world (see Fodor 1981 and 1998). I accept informational atomism without argument.

> the learning of a *new* concept (viz., a concept distinct from Φ). (Fodor 1975, p. 95)

If this is true, then it also works against complex concepts. The individuating condition has to be a single condition. Even if there are different components, there has to be some one thing that assembles them into a condition for that concept. And so, Fodor concludes, "there can't be any such thing as concept learning" (Fodor 2008, p. 139). (Notice that if Fodor really wanted to hold on to the view according to which 'innate' means 'not learned,' he would need to conclude that even the most complex phrasal concepts (such as the concept LIVES IN CHICAGO AND EATS MANGLEWORTS) must be innate. This would make his view even more implausible.)

The conclusion according to which the Standard Argument leads to a radical concept nativism is, however, too quick. All that Fodor's argument shows is that the acquisition of concepts of any kind cannot be accounted for in terms of concept learning. Concepts, then, must be acquired in another way. This other way, for Fodor, consists in concepts being triggered by sensory experience. It is far from obvious, however, why triggering is supposed to require that concepts themselves are innate.

What we are dealing with here seems to be just a terminological matter. Fodor makes the assumption to the effect that either there is a rational connection between a concept and what it represents, or the concept is triggered. And then he simply *defines* the process of triggering as involving some innate items. He says, "you can only trigger a concept that's there, genetically specified, waiting to be triggered" (Fodor 1998, p. 129). But why should we think that a concept can be triggered only if it is already there, in the innate endowment of the organism, waiting to be released by the right kind of stimulus? And also, what would it mean for concepts to *be there, waiting to be triggered*? Unfortunately, nowhere in Fodor's writings do we find a satisfactory explanation of what triggering is really supposed to be or why it requires innate concepts.

The d/D problem

According to Fodor, the crucial feature of the triggering relation between primitive lexical concepts and their occasioning experiences is that it seems random (cf. Fodor 1980, p. 280). Looking at the experience that triggered a concept does not explain why exactly this concept, and not a different one, was acquired. This is what Fodor calls 'the d/D problem': we have no way to explain why it is experience with doorknobs, and not giraffes, that causes the concept DOORKNOB to be acquired. In concept learning, "the organism's knowledge of its environment is exploited to confirm (or disconfirm) generalizations about the extensions of concepts" (Fodor 1975, p. 94, footnote 28). On the other hand, "triggering stimuli may have an *arbitrary* relation to the structures they release" (Fodor 1975, p. 94, footnote 28). In the case of triggering, Fodor emphasizes, the environment in fact provides a *poor* basis for the concepts that the organism acquires (Fodor 1981, p. 280). The data sample from the environment is 'fragmentary and impoverished,' and it is just not sufficient to explain why a given concept is acquired. If the environment does not provide a sufficient explanation of how we acquire concepts, then, Fodor suggests, we should look for an explanation on the side of the organism.

According to Fodor, what *could* solve the d/D problem is exactly the hypothesis testing method, a method which does guarantee that there is a reliable relationship between experiences and concepts that they cause. After all, doorknobs are certainly the best source of evidence for DOORKNOB. Unfortunately, as we have seen, concept learning is of no use when we want to explain acquisition of primitive concepts. In addition, we may question whether hypothesis testing is really immune to the d/D problem. Fodor holds that it remains unexplained why we get the concept DOORKNOB from experiences of doorknobs, and not from experiences of cats. But it seems that he should worry as well about why we form hypotheses about doorknobs, when perceiving doorknobs, and not when perceiving cats. The d/D problem, therefore, arises equally with hypothesis formation and triggering, and it cannot be taken to require concepts to be innate only in the latter case.

Fodor's approach to nativism about concepts seems to be similar to that of Descartes. For Descartes, when we say that external things trigger in us sensory concepts (or sensory 'ideas,' in Descartes's vocabulary), we do not mean that external things through the organs of sense actually transmit concepts into our heads. Rather, he suggests, what we mean is that they transmit "something which gave the mind occasion to form these ideas, by means of an innate faculty, at this time rather than at another." From this Descartes immediately infers that:

> ideas [or concepts] of movements and figures are themselves innate in us. So much the more must the ideas of pain, color, sound and the like be innate, that our mind may, on the occasion of certain corporeal movements envisage these ideas (Descartes, "Notes Against a Certain Program"[7]).

What we see in this passage is a quick switch from an 'innate faculty' or mechanism to innate ideas. Neither Descartes nor Fodor seem to think that such a switch requires any kind of justification. Instead, both assume that if the mind comes up with concepts as a result of being stimulated by external objects, it must be the case that the concepts 'were already there, waiting to be triggered.' Innate mechanisms, however, do not necessarily imply innate concepts. Even though talking about the sensorium does not immediately provide us with an explanation of how sensory concepts are acquired, still, it is far from obvious that the mechanisms which allow us to obtain these concepts also require that concepts themselves be innate.

Innate concept forming mechanisms vs. innate concepts

Even the empiricist idea of concept learning requires some amount of innateness. If we acquire any concepts by means of hypothesis formation and testing, then what is needed to account for that is an innate mechanism, an innate 'machine' for producing inductive inferences (see Fodor 1981, pp. 268–269 for what such a machine must include). In the case of primitive concepts, there also must be some innate *mechanisms* that realize the function from sensory stimuli to sensory concepts. In his 1998, Fodor actually notices that on the account he proposes:

> All that needs to be innate for RED to be acquired is whatever the mechanisms are that determine that red things strike us as they do; which is to say that all that needs to be innate is the sensorium.

And he continues:

> The 'innate sensorium' model suggests that the question how much is innate in concept acquisition can be quite generally dissociated from the question whether any *concepts* are innate. The sensorium is innate by assumption …But …the innateness of the sensorium isn't the innateness

[7] Cited in Adams (1975, pp. 76–77).

> of anything that has intentional content. Since the sensorium isn't an idea, it is a fortiriori not an *innate* idea. So, strictly speaking, the innate sensorium model of the acquisition of RED doesn't require that it, or any other concept, be innate. (Fodor 1998, p. 142)

Fodor clearly does not like this conclusion. Everybody agrees that the sensorium is innate, so what we have here is just a boring claim. Whether concepts are innate should not, however, be decided on *a priori* grounds. Rather, in order to determine what position in the nativism/empiricism debate is the most plausible we need to focus on what exactly we mean by the term innate (see below) and also on what account of concepts we endorse.

3. What is innate in cognition?

The problems with assessing the plausibility of Fodor's mad-dog concept nativism are exacerbated by a general lack of clarity with respect to the meaning of the term 'innate.' One may have the impression that the answer to the empiricism/nativism debate depends merely on what definition of nativism we accept. Say you like the idea of innate concepts. Well, you can go ahead and define 'innate' in such a way that concepts just must be innate. But then, why bother with the whole discussion? Let me clarify at this point that I do *not* intend to argue that Fodor should or should not be a concept nativist in some universal sense of 'innate' because, obviously, there is no such thing. My goal is more modest. What I want to show is that given Fodor's account of concepts and his account of the architecture of the mind, and also given the definition of innateness I support, it does not make sense for Fodor to hold on to his mad-dog concept nativism.

First of all, it should be clear by now that, with respect to the present considerations, we are interested in what can be called *developmental* nativism. We are discussing theories concerning cognitive development, that is, theories about how human beings acquire concepts, develop their language skills, etc. We are *not* interested in issues concerning *epistemological* nativism and empiricism that refer to theories about how knowledge claims are to be justified.

Secondly, within developmental nativism, we distinguish between *representational* nativism and *architectural* nativism. Representational nativism is the view according to which some mental representations (that is, some beliefs, some types of knowledge or some concepts) are innate. On this view, the mind possesses some pre-specified representational content, where different explanations are offered of the innateness or pre-specification of the mental representations in question (I talk about this below). Architectural nativism does not focus on the *output* of the cognitive process, or on the character or source of our representations. Rather, its center of attention is the process of cognition itself. Architectural nativism holds that there is an innate *structure* or an innate *functioning* of our minds. It says that the architecture of the mind is pre-specified. The mind is organized, at birth (or before birth), into some innate structures, and there are some innate ways in which information coming from the external world must be processed by the brain so that the process of cognition can take place. It is commonly assumed as obvious (by both empiricists and nativists) that some sort of architectural innateness is true.[8]

Fodor's mad-dog nativism is an extreme version of *representational* innateness. Our main task at this point is, then, to figure out how representational nativism could be explained in more detail, and what exactly we mean when we talk about innate mental representations. Let us first look at various more general explanations of the term 'innate.'

Meanings of the term 'Innate'

On the most commonly known version of nativism, a feature is considered to be innate if it is possessed from birth (cf. Locke's *Essay Concerning Human Understanding*). There are, however, many features considered to be innate (for instance, secondary sexual characteristics) that are not present at birth. In addition, it is possible that some traits that are present at birth were actually learned in utero (e.g., the newborn's capacity to recognize certain smells or sounds).

'Innate' may also be taken to mean something that is *a priori* to the operation of the senses, something that is not derived from the operation of the senses, but from the mind itself, or more generally, something that

[8] The empiricist *tabula rasa* metaphor is meant to indicate that the mind is originally empty of *objects* of thought (e.g., of ideas, concepts or knowledge). Empiricists do not deny that the mind is endowed with various natural faculties such as perception, understanding or memory, and with certain innate mental powers (e.g., abstraction or comparison).

is completely independent of sensation (cf. Tavuzzi 1987). An innate feature, on this interpretation, is also sometimes defined as "the product of interactions internal to the organism" (Elman 1996, p. 23). Fodor himself hints at this meaning of innateness when he considers whether 'innate' means 'not acquired' (Fodor 2008, p. 132), or, more specifically, 'not acquired in consequence of experience' (Fodor 2008, p. 144), and also when he says that concepts are innate because they do not come from the environment, but rather from the organism (Fodor 1981, p. 280).

This definition of innateness, however, is also not very plausible. If it were accepted, it would turn out that there is really *nothing* innate because basically everything in the human development requires at least minimal interaction with the environment, and thus, also sensory experience. We would be inclined to think, for instance, that walking is an innate capacity of quadruped mammals. But walking is not independent of sensation, of some sort of physical circumstances (e.g., at least gravity and friction are needed), and of some kind of interaction between the organism and the environment (at least, the organism needs to breath, eat and drink in order to be able to ever acquire this innate feature).

Fodor's explanation of innate traits as traits that are acquired by means of the brute-causal process of triggering avoids the problems faced by the first two explanations of innateness. A trait that is acquired by means of triggering does not have to be present at birth. Also, it is quite likely that the process of triggering involves interaction with the environment, and some kind of sensory experience. However, as we have seen, Fodor claims that 'innate' is the only alternative to 'learned,' and that 'learned' is equivalent to '*acquired by means of hypothesis testing.*'[9] Such a claim has some obviously implausible consequences. If something is 'innate' just because it is 'not acquired by means of hypothesis testing,' then too many things (AIDS, for instance) would count as innate. Also, consider the scenario in which a person acquires knowledge of, say, Latin, by swallowing a special Latin-pill (the example comes from Fodor 1975 and 1981). We don't want to say that the knowledge of Latin is innate. For Fodor, however, it would have to be so. The person's knowledge of Latin is not acquired by means of hypothesis

[9] Hypothesis testing is the only *rational cognitive* or *rational psychological* process that Fodor considers as a possible way in which we acquire concepts.

testing, and so, it must have been *triggered* by the pill. If it was triggered, it means that it is innate: it had to already be there, waiting to be triggered.

A possible solution to the Latin pill problem would be to come up with an additional constraint on what it means to be innate. We could say, for instance, that a trait is innate if it satisfies the *normalcy constraint*, that is, if a given organism will acquire it 'in the normal course of events.' Our capacity to walk, then, is innate because it is successfully developed whenever a healthy baby is raised in normal circumstances. On the other hand, our knowledge of calculus is not innate, because it won't be acquired unless special circumstances (learning in a classroom) take place, and it requires the employment of various psychological processes of teaching, learning, reading, studying, etc. The knowledge of Latin from the example above is not innate on this modified version of innateness because it is acquired in a very unusual way.

The concept of 'normalcy,' however, brings with itself its own problems. There may be cases when it is "the very abnormality of the conditions of acquisition that points to the innateness of capacities acquired" (Khalidi 2002, p. 256). So, for instance, if some species of birds are able to "develop adult song even when reared in isolation from conspecifics," then this is "taken as evidence that birdsong in these species is innate" (Khalidi 2002, p. 256). Also, the definition of innateness which refers to the normalcy condition would allow too many traits to count as innate. For instance, it would imply that getting colds is an innate property of humans. Normal conditions can and do change over time. In the modern world, being taught to read is pretty normal. So, because humans now learn to read in the normal course of development, reading would be considered an innate trait.

The conclusion, therefore, is that the fact that a certain trait is always acquired in normal circumstances on its own does not say anything about whether the trait is innate or not. At the same time, it seems that any plausible definition of innateness has to take into consideration some kind of minimally understood 'normal circumstances.' This is because for any living organism there are various requirements that must be met so that it can function properly—without them, no traits, whether innate or not innate, can be acquired.

Consideration of the normal course of events or normal circumstances, as well as the suggestion that those traits count as innate which manifest themselves in all normally developing members of a given population, takes us back to a biological understanding of innateness. This, I think, is the direction that we ought to take in our search for a plausible account of innateness. There are, however, many different senses in which the term ‘innate’ is used in biological sciences. Biologists sometimes consider as innate the following:

- traits that lack malleability (or that *do not* manifest developmental plasticity; traits that are hard to change, because they are insensitive to environmental changes);
- traits that are characteristic of particular species (traits that are typical, universal or exclusive to a species; traits that reflect what it is to be an organism of a given kind);
- traits that are evolutionary adaptations (that are the result of natural selection);
- traits that are genetically determined;
- traits that are unlearned;
- traits that develop in the absence of contact with conspecifics;
- traits (behaviors) that develop fully formed in animals that have been prevented from practicing them (cf. Griffiths 2002, pp. 72–74).

Unfortunately, there is no agreement on how to define these supposed characteristics of innate traits (e.g., it is far from clear how a trait which is genetically determined should be defined). In addition, it turns out that the listed characteristics do not pick out the same traits as innate. A feature may turn out to be innate in one of the senses, but not in another. For instance, a trait may be universal to all members of a species, but learned (e.g., there are birds that have innate song patterns, but in order to trigger these patterns it is required that the birds have contact with other members of the same species; Maclaurin 2002, p. 106). Or it may happen that some trait is a result of natural selection, but does require social interaction for its development (Griffiths 2002, p. 74).

My suggestion is that a plausible definition of innateness should both acknowledge the scientific (i.e., biological) origins of the term ‘innate,’ and enable us to answer *two* questions about traits that are considered. First of all, it has to put limits on *how* an organism can end up possessing an innate trait in question. Secondly, an explanation is needed of *why* the organism has that trait in the first place. Answers to these two questions indicate certain desiderata that any plausible explanation of the term ‘innate’ should satisfy.

First question: How are innate traits acquired?

The answer to the first question will have several parts. First of all, we allow an innate trait to be acquired 'in time.' That is, for any given innate trait, there will always be a time t1 when the organism did not have it and another time t2 when it did. (The time t2 may occur still in the womb, at the moment of birth or at a later stage of development of the organism.) This way, we arrive at the first desideratum:

> Desideratum 1: An innate trait may fail to be present at birth.

Secondly, there will be some processes of acquisition, for instance learning (in the most common sense meaning of the term, as in learning calculus in school) or swallowing a second-language-pill, such that if a trait was acquired by means of them (and could not be acquired without them, excluding science fiction and other extraordinary circumstances), then it is certainly not innate. Acquiring French as a second language will count as an instance of learning; French as a second language is, therefore, not innate. (If French is acquired as a result of being hit in the head, it still does not count as innate.)

> Desideratum 2: An innate trait is not learned.

Third, since *all* features in our phenotype depend basically on two things, our genetic makeup and the influence of environmental factors (it also depends on gene products, such as RNA and protein, and on different interactions between the genes and the environment), it makes sense to say that our innate traits are genetically determined (and genetically inherited) and require certain environmental factors. And so, we have two more desiderata:

> Desideratum 3: The acquisition of an innate trait requires and depends on environmental factors.

> Desideratum 4: An innate trait is a trait that is genetically determined (and genetically transmitted from one generation to another).

Innate traits are somehow coded in the genes (and genetically inherited)[10] so that, in normal circumstances, all normally developing members of a given species are going to possess them (it is not unconditionally guaranteed that the organism will actually acquire a genetically determined trait; serious damage or a disease may prevent acquisition of the trait). In other words, a phenotypic trait is innate for a given genotype if it is sufficiently invariant (Samuels 2002, p. 242), that is to say, if it will emerge even if there are some (not too drastic) changes in the environment, and even if the organism itself undergoes some (again, not too drastic) changes. So, given the normal circumstances (not too rich, as in the calculus class, and not too impoverished, as in the case of a child being raised in complete isolation from other human beings), all the genetically inherited traits that are developed by normal members of a given population would count as innate.

A now obvious objection to this explanation of genetic determination is that it could still allow us to identify too many traits as innate. Most likely all normal adult human beings in a variety of environments have the belief, 'water is wet.' This belief is invariant and it must have some genetic foundations, but we would not want to say that it is innate. Samuels offers the following diagnosis of the problem:

> The fundamental flaw to which all invariance accounts are subject is that they attempt to explain the central features of innateness *solely* in terms of a *mapping* relation between genotypic and phenotypic traits, without imposing any substantive constraints on the mechanisms or processes in virtue of which such mapping relations obtain. What they all ignore, in other words, is the question of what *explains* the existence of such invariant mappings. (Samuels 2002, p. 245)

My suggestion is that what can solve the problem is the explanation of *why/in what sense* a given trait is invariant.

Second question: Why do we have innate traits?

Given that the greatest concern of the present considerations is the phenomenon of the *human* capacity to think by means of concepts, and that

---

[10] There is no agreement with respect to how a trait that is genetically determined should be defined. For more details, cf. Samuels (2004, pp. 137–138; Symons 1992, pp. 140–402).

the main question that we discuss is how concepts are acquired by human thinkers, what we need is a definition of a trait that is innate *in a*—in this case human—*species*. (For this reason, we do not consider as innate those traits that are produced by random mutations in an *individual* organism, and then inherited by its offspring.) In order to answer the question *why* an organism belonging to a given species has an innate trait that it does, we refer, in the end, to evolution. And so, we say:

> Desideratum 5: An innate trait is a result of natural selection; it is an evolutionary adaptation.[11]

To claim that a trait is an adaptation means:

> to make a claim about the *past*: …[it] refers to something produced in the past by natural selection. …It is also to make a claim about *design*. …When one claims that a feature of an organism is an adaptation, 'one is claiming not only that the feature was brought about by differential reproduction among alternative forms, but also that the relative advantage of the feature vis-à-vis its alternatives played a significant causal role in its production.' Finally, given modern understandings of the genetical basis of reproduction, to claim that a trait is an adaptation is to make a certain kind of claim about *genes*. (Symons 1992, p. 140)

'Evolutionary adaptations' are not the same thing as 'adaptive traits.' Adaptations are those traits that are the result of natural selection; they occur through a combination of successive, small, random changes in traits and become fixed in a population by the selection of those variants which have the greatest survival chances. Adaptive traits are those traits which (now) increase the fitness of organisms that possess them. Some of the now adaptive traits, if they do increase survival chances (and if they are heritable), may in the future become fixed for a given population. Until they are fixed, they would not be considered innate. So, what traits are innate will change over time. Supposedly, there was a certain moment in the past when the capacity for language appeared as a result of just some random mutations. After that, this capacity turned out to be an adaptive trait—creatures endowed with language were more likely to reproduce. Now, for the last several thousands

[11] On my account, innate traits are *not* those features which have arisen from a random evolutionary force like statistical drift (cf. Pinker 1997, p. 36). Rather, they were produced by natural selection, which is not random.

of years, language capacity can be considered an evolutionary adaptation,[12] an innate feature of human beings. It is also possible that a trait that is an evolutionary adaptation (and so, innate) currently does not have the same (or any) survival value for the species (it may not be an adaptive trait anymore). For a trait to stop being innate, it must disappear. (It is important to emphasize that whether a trait is an adaptation is an empirical matter. For this reason, we may expect not to know sometimes whether or to what degree a trait is innate.)

At this point, let us see how the proposed definition of innateness can be applied in the case of *representational* nativism. Next, we will clarify what it could mean to say that concepts are innate in the sense of being genetically inherited evolutionary adaptations. Finally, we will assess how plausible such a view is, and whether Fodor is indeed committed to it given his account of concepts and of the architecture of the mind.

4. Kinds of Representational Nativism

What does it mean for mental representations, such as knowledge, beliefs or concepts, to be innate? In order to answer this question we need to distinguish two subcategories within representational innateness: *tissue* and *organism* nativism (cf. Samuels 1998, pp. 559f). *Tissue nativism* is a view according to which to say that there is innate knowledge is to claim that some representations are hard-wired into the brain, and that various neurons are "born 'knowing' what kinds of representations they are destined to take on" (the definition comes from Elman 1996, pp. 26–27). More generally, TN claims that in some cases it is innately specified that various groups of cells (in a normally developing organism) will take on some specific function. For instance, it may be that specific neurons in the cerebral cortex know in advance (perhaps even before they reach their destination when they migrate from the proliferation center, that is, from the place where they are born) that no matter what happens, they can and will be responsible only

[12] This definition of innateness indicates that different traits may be innate to a different degree. The longer a trait is an adaptation, the more innate it is. Language is less innate than, e.g., the functioning of the circulatory system, and other traits that the old parts of the brain are responsible for.

for vision (this is equivalent to saying that on tissue nativism, innate traits lack developmental plasticity).

Organism nativism is not concerned with features of various pieces of brain tissue. Rather, it is a thesis about *whole organisms*. It holds that:

> it is innately specified that organisms possess certain mental representations, … [that is,] there are some types of mental representations (e.g. the type [MOTHER] or RED), such that it's innately specified that we all possess tokens of that type of representation. (Samuels 1998, p. 559)

According to organism nativism, beliefs (or concepts) that are innate for human beings are in some way genetically programmed to arise in our mind; i.e., innate representations are phenotypes that all humans (at least all those with similar enough genotypes, and those who develop in normal conditions) have in common.

Samuels (1998) claims that contemporary theorists who defend nativism about representations, Fodor included, endorse organism nativism; they are "concerned with claims about what innate mental representations *people* (and other organisms) possess and *not* claims about the properties of specific pieces of neural tissue" (Samuels 1998, p. 560). We know that this is the case, Samuels points out, because pretty much everybody agrees that mental representations are *multiply realizable*, that is, that "tokens of the same type of psychological entity (state or process) can be realized by different kinds of neural entity (state or process)" (Samuels 1998, p. 560). And, according to Samuels, tissue representational nativism wants to deny multiple realizability.

Organism nativism, however, could not work as a general definition of nativism. There are many traits that all members of a given species will develop in normal conditions, but these traits do not have to be innate. Their invariance may be due to environmental invariance, they may not have been genetically transmitted and they may have nothing to do with survival value for the species. In addition, recall that the question we keep struggling with is whether *concepts* as thought-parts are innate. If we assume that organism nativism is true and that it is innately specified what concepts we are going to end up with, does this mean that concepts are innate? Well, first of all, if all that organism nativism says is that (1) the organism will acquire concepts (assuming it functions normally) and (2) what kind of concepts it acquires is determined by the kind of world in which it lives and the kind of neurobiology it has (the kind of species it belongs to), then, clearly, Fodor is a nativist, but so is almost everybody else. Organism nativism with respect to various mental representations would be hard to refute. It does not seem very controversial to claim that it is an evolutionary adaptation that while *all* sheep (quickly) acquire the belief 'wolves are dangerous,' all human infants quickly acquire the beliefs 'unsupported objects generally fall,' or 'the parts of the same object generally move together' (cf. Fodor 2000, p. 93), and that ducks have the innate concept MOTHER that they apply to the first moving object that they see right after they have hatched out of an egg. As we shall see below, however, it is not very likely that what this requires is *innate propositional knowledge*, or *innate thought-parts*.

Neural nativism

My suggestion is that if we want to find a place for Fodor's mad-dog nativism in the discussion about representational nativism, we will find it—contrary to Samuels' claim—on the side of tissue nativism. We should notice, however, that there are in fact *two* distinct versions of tissue nativism. Consider the following two ways in which tissue nativism can be summarized:

> I. Various neurons are 'born knowing' what kinds of representations they are destined to take on.
> II. Some representations are hard-wired into the brain; they are encoded in the brain as particular patterns of synaptic connectivity within a specific neural system. (Elman 1996, pp. 26–27)

The first description expresses the kind of tissue nativism presented above. It claims, as we have seen, that as a result of natural selection, specific pieces of tissue are genetically determined to take on certain functions, assuming that all goes well with a given creature's development. The second claim expresses what I call neural nativism, and it defines mental representations as 'fine-grained patterns of cortical activity, which depend on specific patterns of synaptic connectivity' (Elman 1996, p. 364). It says that (at least) some mental representations are hard-wired into the brain; that is, they are *in advance* encoded as particular patterns of synaptic connectivity within a specific neural system (Elman 1996, p. 26) and in specific locations in the brain. According to neural nativism, those pre-specified neural structures are genetically determined to represent specific objects: They have been inherited by the individual and evolved in the species because there is an adaptive value such that the activation of a given neural pattern constitutes thinking a thought containing a specific thought-part, and so constitutes the occurrence of a given concept. Whenever a given neuronal structure fires up, the organism entertains a given representation.

If what neural nativism describes is the *actual* state of affairs regarding the human possession of concepts, then this confirms Fodor's mad-dog nativism. For Fodor's nativism to be plausible, it would have to be the case that it is genetically determined and an evolutionary adaptation that for any kind of stimulus that a (human) cognizer can register, there are certain specific neuronal patterns in specific parts of the brain 'waiting to be triggered'; any cognizable object will (and can only) be represented by some pre-specified neural structure, realized by particular patterns of neural activations in a specific location of the brain. On such an interpretation, Fodor's position turns out to be:

> a type-type identity theory according to which every type of mental entity is identical with some type of neural entity—e.g. that being pain = C-fibre activity, being the concept RED = 30 MHz activity in the frontal cortex and so on. (Samuels 1998, pp. 561–562)

Neural nativism which would support Fodor's mad-dog nativism claims that all primitive concepts that ever appear in our minds are innate, in the sense of being hard-wired. All those patterns of neuronal activity that constitute the thinking of primitive concepts are pre-specified: they are ready to fire prior to experience, they will fire when stimulated by the right kind of a trigger and they are assigned to specific parts of the brain. The exact character of every such pattern that represents any given object is genetically determined and fixed in advance by natural selection, independently of experience.

What would science say?

How plausible, from a scientific point of view, is this kind of radical representational nativism? And what exactly can it mean for our genes to code for innate primitive mental representations? Consider what is called the Hebbian theory of cell interaction, according to which,

> when one cell repeatedly assists in firing another, the axon of the first cell develops synaptic knobs (or enlarges them if they already exist) in contact with the soma of the second cell. (Hebb 1949, p. 63)

As a result,

> any two cells or systems of cells that are repeatedly active at the same time will tend to become 'associated,' so that activity in one facilitates activity in the other. (Hebb 1949, p. 70)

(The theory is often summarized as '*cells that fire together, wire together*.')

For our genes to code for innate mental representations, it would mean that they determine, prior to experience, exactly what cells, in what configurations and in what parts of the brain, need to be excited to arouse a given concept. (If we were to distinguish different concepts at the neural level, we would need to be able to distinguish different patterns of excitations of different neural cells.) In order to make sense, such a view would also require the existence of a mechanism that would guarantee the right connection between a trigger (that is, the object that ends up being represented) and the pre-specified neural pattern 'waiting' in the brain—a neural pattern which, when activated, represents the object. We could say, perhaps, that at some point in the past, this connection used to be established as a result of experience. Perception of a given object would trigger certain cells in a specific location in the brain to start firing together. In agreement with the theory according to which cells that fire together, wire together, a new neural pattern would be formed. In order for mental representations to become innate, it would have to be the case, first of all, that possessing specific kinds of neural patterns is a heritable trait. In addition, it would have to be an adaptive trait for the organism to have it pre-specified independently of experience what concrete patterns of neural activation will stand for any given (primitive) mental representation.

How plausible is such a view? While the issue is far from settled, neural nativism seems *unlikely* in light of current scientific evidence. What scientists can determine is which parts of the brain are (the most) active during various cognitive tasks (cf. Elman 1996, pp. 3–5). They have no way of saying, however, what exactly happens in the brain when I think 'CAT' (cf. Elman 1996, p. 4; Merritt 2008, p. 184) and so there is no certainty whether any kind of *concept* nativism, whether extreme or partial, is the case.

What science does suggest is that specified localization is not sufficient for innateness, and that neither specified localization nor identity of neural patterns across individuals are necessary for innateness. So, first of all, it is not necessarily true that if a function is localized in specific regions of the brain, then it must be innate. Rather, it seems to be the case that various regions of the brain specialize only in virtue of experience.[13] Also, it turns out that the same outcome "can be achieved in a number of ways, i.e., with different forms of cortical representation, and with the collaboration of several different brain regions, in several different working coalitions" (Elman 1996, p. 247). For instance,

> [b]rain organization for language and other higher cognitive functions may vary markedly across normal individuals, in idiosyncratic patterns that are as unique as their finger prints. (Elman 1996, p. 248)

There are no grounds for positing the existence of exactly the same neural patterns in the same parts of the brain that would correspond to the same concepts in different people.

Finally, and most importantly, it is hard to imagine what evolutionary advantage there would be to have pre-specified neural patterns for each mental representation that the organism could entertain. Perhaps it would be plausible to claim that there are some innate (that is, fixed) patterns of synaptic connections in the brain that govern the *functioning* of the heart, of the circulatory system or of visual edge detectors. It would clearly be advantageous if the brain did not have to learn new patterns in order to guarantee the reliability with which the heart and the circulatory system work, and with which neurons in the visual cortex detect edges. Even here, however, the brain must *acquire* these patterns at a certain point in its development.[14]

---

[13] It turns out, for instance, that there is a specific region of the visual cortex that is responsible for spelling, which certainly is not an innate capacity (cf. Elman 1996, p. 242). Similarly, various studies show that there are areas of the cortex "that are uniquely active" in "skilled chess players at different points across the course of the game" (Elman 1996, p. 242). And chess playing is not an innate activity, either.

[14] As Elman points out, "biologically plausible network models have been constructed which demonstrate that such specialized response properties do not have to be prespecified. They emerge naturally and inevitably from cells which are initially uncommitted, simply as a function of a simple learning rule and exposure to stimulation" (Elman 1996, p. 5). So, most likely, what is innate in these cases is that some neural patterns get hard-wired

With respect to conceptual mental representations[15], it seems that the evolutionary advantage is the opposite: to have neurons, or neural networks, that are initially uncommitted, but which, when exposed to stimuli, learn various patterns to encode various representations. This solution may be costly with respect to the amount of time required for the acquisition of mental representations, but it allows for greater flexibility. There is no need to posit (very implausibly) any fixed-in-advance-by-natural-selection patterns of synaptic connections whose role it would be to represent, say, the concepts COMPUTER or RIPSTICK (a new kind of skateboard). Flexibility is what allows us to learn a basically infinite number of different concepts; also, because of it we are able to learn throughout our lives. Because of the brain's flexibility with respect to mental representations, we are able to forget, or to get rid of, useless knowledge in order to make space for new knowledge. Improvement is also possible when some neural patterns get reinforced and other patterns disappear. This also puts a lesser burden on the genome. Genes are responsible for the complexity of the human brain, for the kinds and the number of neurons that we have at birth, for the initial connections between neurons, etc. There is no need for genes to also code for a specific neural pattern for every possible concept.[16]

At the brain level, to acquire an innate trait means to acquire a certain more or less complex set of neural patterns that is an evolutionary adaptation. In those cases where the possession of these specific patterns of synaptic connections is not an evolutionary adaptation, the traits that they are responsible for are not innate. (When I learn how to rollerblade and new patterns of synaptic connections are produced in my brain, the exact character of these patterns is certainly not pre-determined by natural selection.)

What seems more likely is that it is *the capacity to acquire concepts* which is an evolutionary adaptation; but how particular concepts are realized

---

very early in the development of the organism, and that they have a high level of priority in the allocation of neural resources.

[15] I do not intend to take a stand concerning perceptual mental representations. It may well be that the basic representations involved in perception, such as edge and motion detectors, are evolutionary adaptations. If that is the case, nativism about basic perceptual categories could follow.

[16] Cf. Elman (1996, p. 8): "There is simply too much plasticity in the development of higher organisms to ignore the critical effect of experience. …there aren't enough genes to encode the final form directly; …genes don't need to code everything."

in the brain is not. The exact character of the new pattern that will stand for a given representation depends on various factors: the person's genetic endowment, the stage in her development (or her age) at the moment when she acquires a given concept, the state of her brain at the moment of stimulation and a myriad of other factors from both external and internal environments. The new pattern most likely will not remain unchanged throughout the life of a given cognizer. If it is successful (if it turns out to be important and advantageous), it will get reinforced; if it is not successful, it will weaken, and the neurons involved in its occurrence will get hired for a new job.

5. Conclusion

Among the various explanations of innateness that we have seen, it is neural nativism that provides a plausible and at the same time substantial explanation of what it could mean for concepts as thought-parts to be innate. Neural nativism, however, is not plausible from a scientific point of view. Given the current state of scientific research, we can also conclude that concepts understood as thought-parts are not innate. The remaining task is to show that Fodor himself has actually no reason to hold on to his mad-dog nativism.

Implications of Fodor's account of the architecture of the mind

The mind, according to Fodor, has three components—three kinds of cognitive systems: transducers, input systems and central processors.[17] Transducers are mental mechanisms that link the cognizing subject to the external world. They register information from the external environment (they receive physical, non-symbolic input such as light, sound, scent, etc.), and transform it into a format that can be understood and used by input systems, i.e., into neural signals, pieces of brain-code, or some patterns of nerve impulses in the brain (cf. Fodor 1983, p. 41). While the role of transducers is to specify "the distribution of stimulations at the 'surfaces'

[17] The main role of central processors is the fixation of belief. Central processors are not relevant for the issue of concept acquisition and so they may be ignored. For a detailed presentation of Fodor's account of the architecture of the mind, see his 1983.

(as it were) of the organism, input systems deliver representations that are most naturally interpreted as characterizing the arrangement of *things in the world*" (Fodor 1983, p. 42); they allow the subject to identify objects in its environment. On the basis of the information provided in symbolic form by the transducers, input systems—by means of computation—infer how the external world really is. As Fodor explains,

> [i]nput analyzers are inference-performing systems. …Specifically, the inferences at issue have as their 'premises' transduced representations of proximal stimulus configurations, and as their 'conclusions' representations of the character and distribution of distal objects. (Fodor 1983, p. 42)

Together with transducers, input systems are responsible for "the realization of a function from stimuli onto primitive concepts" (Fodor 1981, p. 265).

On Fodor's account, sensory experience, external stimulation and properly functioning transducers are necessary for the acquisition of sensory concepts. These are *architectural* constraints on sensation. Of course, the fact that transducers require external stimulation does not yet imply that (sensory) concepts are *not* innate. It may well be the case, and Fodor thinks it is the case, that the output produced by transducers only *triggers* concepts that are 'already there.' On the other hand, Fodor's account of transducers does not indicate any need for innate concepts. Concepts themselves do not yet enter the stage at this point in the process of cognition. All that the theory of transducers says is that there must be *mechanisms* which will transform physical stimulation from the external world into neural signals. The character of a particular pattern of neural activations that appears at the output of transducers does not have to be pre-specified. It is rather determined (mostly) by the pattern or shape of the sensory excitation that causes it. And so, while Fodor's doctrine of transducers implies architectural nativism, it neither obviously requires nor precludes representational nativism.

We find a more substantial nativist commitment in Fodor's treatment of input systems. Each input system, for Fodor, is associated with a fixed neural architecture: "hardwired connections …facilitate the flow of information from one neural structure to another" (Fodor 1983, p. 98f). This is, again, architectural nativism. But then, Fodor also talks about input systems having 'knowledge' and 'beliefs,' and 'making inferences.' Each input system, he says, "comes with a database which is, in effect, what it innately believes about its proprietary computational domain" (Fodor 2000, p. 1); "much of the information at the disposal of [input] systems is innately specified" (Fodor 1983, pp. 100–101). And so, Fodor states:

> qua modularity theory, the kind of nativism we're imagining …postulates features of innate cognitive *content* as well as features of innate cognitive architecture. (Fodor 2000, p. 91)

Input systems, according to Fodor, are useful in contributing to a creature's fitness because the world in the end shapes what the mind believes. The beliefs that the mind holds to be true are shaped by experience; that is, they are formed by processes "that are sensitive to the way the world contingently is" (Fodor 2000, p. 95). If some among the beliefs that the mind holds were to be innate, they could be of any use only if they were true. And for this they would have to also have been shaped by experience. Based on this reasoning, Fodor infers that it only makes sense to assume that the innate beliefs in the input systems' databases have been produced by natural selection (Fodor 2000, p. 94). Current innate beliefs in the input systems' databases, according to Fodor, were previously acquired by experience. At some point, it turned out that those members of our species who had those beliefs were more likely to live long enough to reproduce. As a result, the holders of the right kind of innate beliefs became the majority. Eventually, the possession of those beliefs was fixed as an evolutionary adaptation.[18]

Still, nothing in the evidence Fodor cites supports his mad-dog representational nativism over an architectural nativism according to which input systems are innately disposed to produce certain representations given particular experiences. Fodor claims that there must be an innate database of input systems, but this may as well be understood as *organism nativism*, which is more architectural than representational. It is the kind of architecture of the mind that we have, plus the world in which we end up living, that together determine what 'beliefs' are needed for the proper functioning of mental modules. This does not require that concepts themselves are innate. Rather, it just seems irrelevant to Fodor's argument for concept nativism.

---

[18] Fodor considers an example of an innate module for avoiding visual cliffs. The module can function properly only because it has in its database an innate and contingent belief that there is a contingent regularity between differences of depth in the actual world and differences of visual texture (cf. Fodor 2000, p. 92).

Consider also Fodor's account of reflexes. For Fodor, reflexes are a good example of an innate faculty. Their innateness can be explained in terms of how they know *what proximal stimulus to respond to and what proximal response to make to it*. If innate knowledge of modules is supposed to be similar to innate knowledge of reflexes, then there seems to be no reason to believe that *concepts* are an innate element of human cognition. There is certainly no propositional knowledge and no thought-parts to which reflexes have access. Their 'knowing,' therefore, is not literal.

In his 1983, Fodor considers himself to be a faculty psychologist of the kind that holds that what is innate in human cognition are *cognitive mechanisms* of some sort. He clearly distinguishes his position from Chomsky, whom he classifies as an advocate of *neocartesianism*. Neocartesianism, Fodor explains, is a view according to which what is innate is "a certain *body of information*" (Fodor 1983, p. 4), or certain *truths* that "human beings innately grasp" (Fodor 1983, p. 7). This view, according to Fodor, is not plausible. He says:

> It may be that the development of arms and the development of anaphora each critically involves the exploitation of a specific genetic endowment. And it may also be that what is innate can, in each case, be described as 'information' in the relatively uninteresting statistical sense that implies only nonrandomness. But there is, surely, no reason to suppose that the development of arms requires access to innately given *propositional contents*. There is nothing that growing arms requires one to cognize, innately or otherwise. (Fodor 1983, pp. 5–56)

According to Fodor, it is only the architectural kind of nativism that Chomsky's views imply. Nothing suggests the need for innate concepts or innate representations. What we need, Fodor explains, is not some kind of "innately cognized rule," but rather "a psychological mechanism—a piece of hardware …whose structure somehow imposes limitations upon its capacities" (Fodor 1983, p. 8). Fodor's proposal against Chomsky is, therefore, to turn to:

> a different notion of mental structure, one according to which a psychological faculty is par excellence a sort of mechanism. Neocartesians individuate faculties by reference to their typical propositional contents (so that, for example, the putative language organ is so identified in virtue of the information about linguistic universals

> that it contains). By contrast, according to the present account, a faculty is individuated *by reference to its typical effects*, which is to say that it is functionally individuated. (Fodor 1983, p. 10)

A similar reasoning, I believe, could be applied to Fodor's own views on the innate databases of input systems, and more generally to his account of concept acquisition. As we said above, the account of both transducers and input systems certainly requires various innate architectural constraints. These cognitive faculties, like reflexes, have to be programmed to respond in specific ways to specific stimuli. This is a constraint regarding how they function, and not what output they produce. It just does not make sense to claim that these faculties are able to do their job only because what they produce as their output are innate mental representations. We do not know how exactly the mind acquires concepts, but this mystery does not imply that the representations that show up as the final product of the process of cognition have to be innate.

Fodor's input systems, as we said, are supposed to be 'computational mechanisms' whose role is to produce representations of distal objects. And so, Fodor explains, visual input systems, for instance, are innately programmed to apply the right kinds of algorithms to the data that they receive from the visual transducers. As a result, they produce visual representations of distal objects. If what we want to consider is the issue of concept acquisition, it seems that what we need to posit are also certain innate mechanisms that apply the right kinds of algorithms to the data received and produce mental symbols representing distal objects. It is plausible that it is simply the *functioning* of our cognitive modules that is an evolutionary adaptation, an innate trait, fixed by natural selection because of its survival advantage. It is advantageous for the human species to have fast and automatic mechanisms, that is, input systems, which produce (usually in a reliable way) representations of distal objects. The representations themselves need not be innate.

Implications of Fodor's Account of Concepts

Similarly, as with the case of Fodor's views on of the architecture of the mind, nothing in his account of concepts as thought-parts indicates that he is justified in supporting his mad-dog concept nativism. In virtue of his physicalism, Fodor thinks that concepts, which are parts of sentences of the language of thought, must be physically embodied; they must be either identical to or constituted by states of the brain. Each different concept must correspond to a distinct state of the brain, to a distinct pattern of neural activation that encodes it. Concepts, therefore, are patterns of neuronal activity; they are symbols of the brain-code.

Given that each concept is a distinct pattern of the neural code, the only way to make sense of Fodor's idea that concepts are innate in the sense of *being already there, waiting to be triggered* is, as we have seen, to accept neural nativism. What Fodor would need is not only the kind of tissue nativism which negates multiple realizability, but also the kind of nativism according to which it is somehow predestined, programmed in the genes, that a given piece of the brain-code *represents* a given feature of the world. This is what would allow Fodor to claim, no matter how implausibly, that we have, for instance, an innate concept RIPSTICK, because there is in our brain some *particular* symbol, some particular neural pattern, produced by natural selection and destined to be a RIPSTICK symbol.

We have seen above that this is not plausible: identity between individual concepts and specific neural patterns in specific locations in the brain is not an evolutionary adaptation. What I suggest is that Fodor himself has no reason to support such a position.

If Fodor accepts the definition of innateness that I propose, he would only need to conclude that it is innate for humans—because it is a genetically determined evolutionary adaptation—to have a very complex brain, with a huge number of neurons, and a great potential for acquiring, storing and modifying new connections and activation patterns. He would not need to say that we have an innate RIPSTICK concept, an innately determined symbol of brain-code. Rather, he could hold a more plausible view according to which we are innately disposed to enlist *some or other* symbol of the brain-code to serve as a ripstick indicator (cf. Prinz 2004, p. 230). When I see a ripstick for the first time, a complex cognitive process gets activated. My sensory organs inform me about certain colors, shapes, textures, noises, etc.; my memory brings up images of skateboards and surfboards, and my past thoughts about picking up snowboarding one day; my language faculty assigns a new word to the newly cognized object. A new neural pattern is produced in my brain. The pattern does not remain unchanged through time. It gets modified the better I become at riding the ripstick. It also changes when I alternate between riding the ripstick, rollerblading and riding a regular skateboard. What is innate in all this is the complex functioning of various cognitive mechanisms.

There is no reason why Fodor should not agree that in addition to various innate architectural constraints (constraints on various cognitive mechanisms, the structure and functioning of sensory organs, etc.), evolution also endowed us with general-purpose detecting and tracking abilities. Because of these innate abilities, new patterns of neural activation (new symbols) are produced in our brains when we acquire a new concept. We do not have to be born with pre-specified symbols of the brain-code. It is enough that we have an innate capacity to 'hire' a neural pattern in response to a given kind of stimulus. We are a successful species because "perceiving objects in our environment" gives us "the concepts that enable us to think about them, and consequently to form beliefs and desires about them" (Davis 2003, p. 456).

References

Adams, R. M. 1975. "Where do Our Ideas come from?" In: S. Stich (ed.), *Innate Ideas*. Berkeley, CA: University of California Press, 71–87.

Cowie, F. 1998. *What's Within? Nativism Reconsidered*. Oxford: Oxford University Press.

Davis, W. 2003. *Meaning, Expression and Thought*. Cambridge, UK; New York: Cambridge University Press.

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

Fodor, J. A. 1975. *The Language of Thought*. New York: Thomas Y. Crowell.

_______ 1981. "The Present Status of the Innateness Controversy." In: J. Fodor, *RePresentations*. Cambridge, MA: MIT Press, 283–292.

_______ 1983. *The Modularity of Mind*. Cambridge, MA: MIT Press.

_______ 1984. "Observation Reconsidered." *Philosophy of Science* 51, 23–43, reprinted in Fodor (1990), 231–251.

_______ 1990. *Theory of Content and Other Essays*. Cambridge, MA: MIT Press.

_______ 1998. *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.

_______ 2000. *The Mind Doesn't Work That Way*. Cambridge, MA: MIT Press.

_______ 2008. *LOT2: The Language of Thought Revisited*. Oxford: Oxford University Press.

Griffiths, P. 2002. "What is Innateness?" *The Monist*, 85 (1), 70–85.

Hebb, D. O. 1949. *The Organization of Behavior*. New York: Wiley.

Jackendoff, R. 1989. "What is a Concept, that a Person May Grasp It?" *Mind & Language*, 4 (1–2), 68–102.

Khalidi, M. A. 2002. "Nature and Nurture in Cognition." *British Journal for the Philosophy of Science*, 53, 251–272.

Locke, J. 1975. *Essay Concerning Human Understanding*. Oxford: Clarendon Press.

Maclaurin, J. 2002. "The Resurrection of Innateness." *The Monist*, 85 (1), 105–131.

Merritt, M. 2008. "Nativism and Neurobiology: Representations, Representing, and the Continuum of Cognition." *Review of General Psychology*, 12 (2), 181–191.

Pinker, S. 1997. *How the Mind Works*. New York: W.W. Norton & Company.

Prinz, J. 2004. *Furnishing the Mind: Concepts and Their Perceptual Basis*. Cambridge, MA: MIT Press.

Samuels, R. 2004. "Innateness in Cognitive Science." *Trends in Cognitive Sciences*, 8 (3), 136–141.

_______ 2002. "Nativism in Cognitive Science." *Mind and Language*, 17 (3), 233–265.

_______ 1998. What brains won't tell us about the mind. *Mind and Language*, 13 (4), 548–570.

Symons, D. 1992. "On the Use and Misuse of Darwinism in the Study of Human Behavior." In: J. H. Barkow, L. Cosmides, & J. Tooby (eds.). *The Adapted Mind*. Oxford; New York: Oxford University Press, 137–159.

Tavuzzi, M. 1987. "Aquinas on the Preliminary Grasp of Being." *The Thomist*, 51, 555–574.