

FODOR AND AQUINAS:

THE ARCHITECTURE OF THE MIND
AND THE NATURE OF CONCEPT ACQUISITION

A Dissertation
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment for the requirements for the
degree of
Doctor of Philosophy
in Philosophy

By

Justyna M. Japola, M.A.
Georgetown University, Washington DC
December 2009

UMI Number: 3390016

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.

UMI[®]

Dissertation Publishing

UMI 3390016

Copyright 2010 by ProQuest LLC.

All rights reserved. This edition of the work is protected against unauthorized copying under Title 17, United States Code.

ProQuest[®]

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Copyright 2009 by Justyna M. Japola
All Rights Reserved

PREVIEW

FODOR AND AQUINAS:
THE ARCHITECTURE OF THE MIND AND THE NATURE OF CONCEPT
ACQUISITION

Justyna M. Japola, M.A.

Thesis Advisor: Wayne A. Davis, Ph. D.

ABSTRACT

Fodor (1975 and 1981b) explains the paradigm empiricist method of concept acquisition as consisting in forming and testing hypotheses about objects that fall under a concept. This method, he notices, can only work for complex concepts, because we have to possess some concepts in order to form hypotheses. If so, then none of our simple (or primitive) concepts can be learned. If we still have them then they must be innate. Aquinas, on the other hand, is famous for his opposition to Platonic nativism, and is universally considered an empiricist with respect to cognition. In my dissertation I show that Fodor's and Aquinas's accounts of the architecture of the mind are quite similar. I argue that because one's position in the empiricism-nativism debate should be a function of one's account of the architecture of the mind, Fodor and Aquinas should be on the same side of the debate. My claim is that they should be on the side of nativism, but not the kind of radical concept nativism that Fodor is famous for. I attempt to show that it is Aquinas who is closer to a successful account of cognition with the required *amount of* and the right *kind of* innate elements. In the end, I aim to show how Aquinas could help Fodor to arrive at a more plausible account of concept acquisition.

TABLE OF CONTENTS

Abbreviations.....	1
Introduction.....	2
Chapter I: Traditional Interpretation of Aquinas as an Empiricist and its Failure.....	18
1.1 The Aristotelian-Thomistic account of cognition.....	18
1.2 Failure of the traditional interpretation of Aquinas as an empiricist.....	37
Chapter II: Innate Elements in Aquinas: the Level of the Senses.....	97
2.1 The need for innate elements in Aquinas's account of cognition.....	97
2.2 The Form Trans-Formation Interpretation – A First Approximation.....	103
2.3 The role of the internal senses and instinct.....	114
Chapter III: Innateness in Aquinas: the Level of the Intellect.....	151
3.1 Aquinas on concepts: Intelligible species vs. mental word.....	151
3.2 The Transcendentals.....	167
3.3 Collaboration of the agent intellect and the cogitative power.....	216

Chapter IV: Fodor: A Contemporary Nativist.....	244
4.1 Fodor's Computational Theory of the Mind.....	244
4.2 Fodor's Account of Concepts.....	252
4.3 Concept acquisition.....	258
4.4 Fodor's views on the architecture of the mind.....	281
Chapter V: What is Innate in Cognition?.....	292
5.1 Most common accounts of innateness.....	292
5.2 Two questions: How and why a trait is acquired.....	306
5.3 Representational innateness.....	321
VI. Conclusion.....	339
Bibliography.....	354

ABBREVIATIONS

Thomas Aquinas

<i>CT</i>	<i>Compendium Theologiae</i>
<i>DEEE</i>	<i>De Ente et Essentia</i>
<i>DP</i>	<i>De Potentia</i>
<i>DQO</i>	<i>De quattuor oppositis</i>
<i>DSC</i>	<i>De spiritualibus creaturis</i>
<i>DV</i>	<i>De Veritate</i>
<i>InBoeth</i>	<i>Commentary on Boethius's De Trinitate</i>
<i>InDA</i>	<i>Commentary on Aristotle's De Anima</i>
<i>InDI</i>	<i>Commentary on Aristotle's De Interpretatione</i>
<i>InDMR</i>	<i>Commentary on Aristotle's De Memoria et Reminiscentia</i>
<i>InIoan</i>	<i>Commentary on the Gospel of St. John</i>
<i>InMetaph</i>	<i>Commentary on Aristotle's Metaphysics</i>
<i>InPhys</i>	<i>Commentary on Aristotle's Physics</i>
<i>InSent</i>	<i>Commentary on Peter Lombard's Sentences</i>
<i>QDA</i>	<i>Questions on the Soul</i>
<i>SCG</i>	<i>Summa Contra Gentiles</i>
<i>ST</i>	<i>Summa Theologiae</i>

Aristotle

<i>DA</i>	<i>De Anima</i>
<i>Metaph</i>	<i>Metaphysics</i>
<i>PA</i>	<i>Posterior Analytics</i>

INTRODUCTION

The main question that I consider in my dissertation is whether the human cognitive endowment is innate, and in particular, whether we have innate concepts. The way I proceed in order to find an answer to these questions is by analyzing the views of Jerry Fodor and Thomas Aquinas, two philosophers who stand on opposite sides of the debate.

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 believes (Fodor, 1975 and 1981b) that the paradigm empiricist method of concept acquisition that he calls 'concept learning' proceeds *via* forming and testing hypotheses about objects that fall under a concept. As it turns out, however, this method can work at most for complex concepts (those that can be decomposed into constituent concepts), because we cannot form hypotheses unless we already have some concepts. This implies that all of our simple or primitive concepts cannot be learned, and since we do acquire them, they have to be innate. Since, as Fodor argues, it turns out that no new (primitive) concepts can be learned, we must conclude that all (primitive) concepts are innate: as he puts it, they are already there, genetically specified, waiting to be triggered. Fodor, then, is famous for being a mad-dog concept

nativist. Aquinas, on the other hand, as seems to be appropriate for a Dumb Ox,¹ is an empiricist with respect to cognition. He is also famous for his wholehearted opposition to Platonic nativism (cf. *ST I 79, 3*; *SCG 2, 77*).²

My suggestion is that one's position in the empiricism-nativism debate should be a function of one's account of the architecture of the mind. As it turns out, Fodor's and Aquinas's accounts of the architecture of the mind are quite similar. If so, then it seems that both Fodor and Aquinas should be on the same side of the debate. My claim is that they should be on the side of nativism, but not exactly the kind of nativism that Fodor endorses.³ Indeed, I attempt to show that it is Aquinas who is closer to a successful account of cognition with the required amount and the right kind of innate elements. Given that there is no universally accepted definition of nativism, my goal in this dissertation is only to show that Fodor should be as much of a concept nativist (or as much of a concept empiricist) as should be Aquinas. In the end, my hope is to show how Aquinas could help Fodor arrive at a more plausible account of concept acquisition.

¹ This is the nickname that Thomas had during his student years. He was huge, and rather silent, and for these two reasons he was often mocked by his classmate. As the story goes, one day Aquinas's teacher, Albert the Great, said in his defense: 'You may call this man a dumb ox but it is his bellowing that will be heard across the universe.'

² See p. 5 for the list of Abbreviations used in this dissertation.

³ At least not the kind of radical concept nativism for which he is most famous (Fodor, 1975 and 1980). It could still be compatible with the kind of nativism which I think is implied by his *Modularity of Mind* (1983).

i. Contemporary Interest?

In the first three chapters of the dissertation I discuss the problem of the architecture of the mind and the nature of concept acquisition in the work of Thomas Aquinas.

Aquinas's main claims with respect to cognition are as follows:

- C1. The only objects that we encounter in the world are *particulars*. Each individual object belonging to a natural kind is an instance of a *substance*, that is, it is a composite of *substantial form* and *matter*.
- C2. There are two main levels of cognition: the senses and the intellect, with the former further divided into two sublevels: the level of the external and internal senses.
- C3. Cognition consists of acquisition of the form of the object cognized by the cognizing subject.
- C4. In the first stage of cognition, an object becomes actually known when a *sensible species*, that is, the object's sensible form, is received by the external senses of the knower.
- C5. External sensation concerns what is particular and material; its vehicles of cognition, i.e., *sensible species*, are cognitive forms that represent accidental features of things.
- C6. On the basis of the data provided by the external senses, the internal senses produce their own vehicles of cognition, called *phantasms*. *Phantasms*, like *sensible species*, represent accidental features of things.
- C7. In the second stage of cognition, the intellect takes the *phantasm* produced by the internal senses and abstracts from it a universal concept (called an '*intelligible species*,' or a '*mental word*').
- C8. Concepts are immaterial and universal forms representing essences of things.
- C9. Even though the proper objects of the intellect are essences, the intellect somehow acquires its objects of cognition from the senses (cognitive empiricism).

C10. The first object of the intellect is Being⁴ ("The first thing conceived by the intellect is being; because everything is knowable only inasmuch as it is in actuality. Hence, being is the proper object of the intellect;" *ST I 5, 2*).

Aquinas's main claims with respect to cognition, when they are expressed in typical scholastic terminology, sound rather unintelligible to the contemporary reader. One of the goals of this dissertation, therefore, is to present Aquinas's account of cognition in such a way that it not only becomes understandable in the XXI century, but also becomes obvious that in Aquinas's texts we find topics, and suggested solutions to problems that continue to trouble and to fascinate philosophers working in the philosophy of mind in our times.

In addition to the question of the possible points of interest in Aquinas for the contemporary reader, I consider four main problems, and two sub-problems that threaten Aquinas's views on cognition. The first two problems concern the objects of the intellectual cognition. The second two problems are related to the issue of the relationship between the sensory and the intellectual levels of cognition. The fifth problem concerns knowledge of individuals and the sixth the transcendental concepts of the intellect.

⁴ I capitalize the first letter of the word 'Being' in all cases where it refers to the crucial notion in Aquinas's philosophy, the notion of Being-as-such, or Being-as-Being (what I call Being in the B-B sense; see below, 3.2 and 3.3). Whenever the entire word is capitalized (e.g., BEING), it refers to the corresponding concept.

ii. Six problems with Aquinas's account of cognition

P1. The first problem concerns Aquinas's view of essences being the proper object of the intellect. As it seems, our thinking in most cases does *not* consist in grasping what constitutes the essence of things. So, the idea that the proper object of the intellect is essence seems to require too much from intellectual cognition. I call it the *Concepts as Essences Problem*.

P2. The second problem, what I call the *Being as the First Intelligible Problem*, is related to Aquinas's claim to the effect that Being (*Ens*) is the first intelligible and that without Being nothing can be apprehended by the intellect. This claim, first of all, sounds very mysterious. Also, given that on Aquinas's account essences in the mind are always universal in the sense that numerically distinct things can have the same essence, and Being always refers to what is individual (everything is a Being, or is intelligible, in virtue of its own unique act of transcendental existence), we seem to have a contradiction: it cannot be that the proper objects of the intellect are both *essences* and *Being*.

P3. What constitutes the most challenging problem in Aquinas's views on cognition is the issue of concept acquisition. It is not clear how sensible species are supposed to become a universal concept, or how the material impression in the sense organ can be transformed into a component of thought. This is what today is called the *Transduction Problem*.

P4. Finally, Aquinas's account of cognition also faces what Jerry Fodor calls the doorknob/DOORKNOB problem, or the d/D problem. The d/D problem concerns the relationship between sensory experiences and concepts produced in our mind. This relationship seems to be entirely random: there seems to be no way to explain why having *those* experiences leads to possession of *this* concept.

The two sub-problems that I also consider are as follows:

P5. The main ontological category for Aquinas is what he calls a primary substance. Since the intellect only cognizes essences of things, and the senses only accidents – it seems that he has no account at all of the cognition of primary substances.

P6. Transcendentals, on Aquinas's account, are the most general concepts which express the most general features of everything that exists. Because of their scope, transcendentals are not very informative. What is not clear is the relationship between transcendental concepts and such concepts as, say, the concept DOG.

iii. The Form Transmission interpretation: Aquinas as an empiricist

The most common interpretation of Aquinas's views, that I call the Form Transmission Account (FT) focuses on the claim C3 to the effect that, for Aquinas, cognition consists in the acquisition of the very form of the object cognized by the cognitive faculties of the cognizing subject. The apparent advantage of this interpretation is that it would avoid

the Transduction and the d/D problems listed above (P3 and P4). If what is in the intellect is the same as what is in the object and what is in the external and internal senses, then the connection between all the levels of cognition should not be problematic.

The FT, however, turns out to be implausible for various reasons. What is most important is that the main claim of the FT, that is, the idea that in cognition *the very form* (numerically the same, or type-identical) of the object cognized is received in the cognitive faculties of the subject, turns out to be just plain wrong once we analyze various passages from Aquinas's *De Veritate* and from his commentary on Aristotle's *De Anima*.

iv. The Form Trans-Formation interpretation: Nativism in Aquinas

The Form Trans-Formation account of cognition (FTF) that I propose takes seriously the claim that cognition consists in the acquisition of form (claim 3). At the same time, the FTF also focuses on claims C5 to C8. If the senses cognize only the accidental features of things (C5), and the intellect – only the essences (C8), then the connection between the two levels is far from obvious. The forms that the different cognitive faculties operate on *cannot* be the same.

The FTF, therefore, admits the seriousness of the transduction problem. It suggests that some kind of *nativism* is needed in order to solve it. This nativism, the FTF notices, is already indicated in Aquinas's own texts. In particular, the FTF is going to

claim that a solution to the transduction problem can be found in Aquinas's own theory of transcendentals and in his account of the process of concept acquisition in terms of the intellect collaborating with the internal sense of cogitative power.

The success of the solution offered by the FTF account of cognition depends to a great extent on the plausibility of two ideas:

- (A) that for Aquinas, only intelligible species should be understood as concepts, and that concepts are thought-parts; and
- (B) that for a trait to be innate it must be an evolutionary adaptation, genetically inherited in a species.

v. What are Concepts?

In chapter III I argue that Aquinas's suggestion that the proper object of the intellect are universal essences *is not supposed to* mean that intellectual cognition always requires a grasp of a rich conception of a thing's essence. On my interpretation, it is still the ultimate goal of human cognitive activities to understand the essence of the thing cognized. However, we should also distinguish a lower level of intellectual cognition, the level of thinking, where we form concepts as *thought-parts* in order to be able to think about things whose essences we (yet) do not grasp.

On my proposal, it is only Aquinas's intelligible *species* that should be understood as concepts in the sense of thought-parts. What he calls 'mental words' are interpreted as *conceptions*, expressing the essences of things. A concept as a thought-part does not have a rich informational content. A (simple) concept does not have a structure.

It is merely a *sign* pointing in two directions: on the one hand, it points towards an individual object that belongs to its extension (and that caused its occurrence); on the other hand, it points towards the object's essence, that is, towards the definition of the *kind* to which the thing belongs.

The idea that for Aquinas concepts are thought-parts serves as part of the reply to the transduction problem concerning the passage from the senses to the intellect. It is much more mysterious to explain the passage from sensible *species* to concepts if the latter require a grasp of the definitions of things. If concepts are thought-parts, on the other hand, i.e., if they are only signs of essences, the task is less challenging. The other part of the solution of the transduction problem is to be found in Aquinas's theory of the collaboration between the intellect and the cogitative power.

First of all, I show that Aquinas's claim that being is the first intelligible is a claim about the transcendental Being. It expresses not a view about *concepts*, but rather a view about *how the mind works*: it is a being-detecting mechanism. Being, together with other transcendentals, are innate mechanisms, innate rules of functioning of the cogitative power.

The cogitative power is to be understood as a sophisticated innate cognitive mechanism. It transforms the information provided by the lower sensorium into phantasms representing individual substances recognized as substances of a certain kind. To these phantasms the cogitative power applies the intellect's concepts. The phantasms at this level of cognition are in fact applications of the intellect's universal concepts to individual things.

Also, we can now see a point of similarity between Fodor and Aquinas. For Fodor, concepts are also thought-parts. More precisely, concepts, for Fodor, are mental entities, parts of propositional attitudes. For a mind to have a concept with a specific content is to have a mental representation with some kind of world-to-symbol causal connection. Because of his physicalism, Fodor also holds that concepts are physically embodied. Each different concept is constituted by 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, as I also show, the accounts of the architecture of the mind of Aquinas and Fodor are considerably similar, it seems that they should reach the same conclusion with respect to the problem of acquisition of our concepts and the empiricism/nativism debate.

vi. Definition of innateness

The question of whether Aquinas and Fodor should adopt the same position with respect to the empiricism/nativism debate about concepts cannot be answered until we figure out a definition of innateness that both philosophers would accept. This task is far from being easy. The problem of innateness is widely discussed in our times, and it has been discussed for centuries. One thing that is certain about it, however, is that there is no one definition of innateness that all philosophers would accept.

An innate trait used to be explained as a trait that is present at birth, or that is acquired by a creature independently of sensation. But today we know that some traits that are considered innate are *not* present at birth, and that it is possible for something to be innate even though it requires sensory experience. Fodor's suggestion that a feature is innate if it is triggered, that is, acquired brute-causally, is also problematic. Fodor contrasts triggering to the rational-causal way of acquisition, and the latter, he believes, is equivalent to 'learned *by means of hypothesis testing*.' Hypothesis testing method, however, turns out to be useless with respect to the issue of concept acquisition. In addition, Fodor's definition of innateness has other quite implausible consequences (for instance, it would classify the knowledge of Latin acquired by means of swallowing a special Latin-pill as innate). Another proposal is to define an innate trait as a trait that will develop in a given species 'in normal circumstances.' This definition, again, will have implausible consequences. Also, 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. The term 'innate,' which took its origins in biology, is also used with many different senses in biological sciences. It may refer to traits that are typical for a given species, genetically determined, inherited, insensitive to environmental changes, etc.

Even though there is no agreement with respect to the meaning of innateness in biological sciences, still it seems that it is a good idea to look for a biological definition of the concept. And so, on my proposal those traits are innate that are genetically inherited and that are evolutionary adaptations; they were produced by natural selection, and fixed in a given population because of their survival advantage for a given species. On

this proposal, if we manage to explain in what sense a trait is innate, we explain not only *how* an organism ends up having the trait in question (how it was acquired), but also *why* the organism has that trait in the first place.

vii. Aquinas and Fodor: The outcome of the debate

Fodor's mad-dog concept nativism according to which (most of) our concepts are innate, that is are 'there, waiting to be triggered,' is a version of *representational nativism*. This view about innateness of some mental representations is opposed to architectural nativism—concerning various mental structures. There is not much disagreement concerning architectural nativism in the domain of cognition. Even if concepts are acquired from experience, there must be some innate structures, more or less developed, for concept acquisition to take place. And so, both cognitive empiricists and nativists agree that we have to posit innate elements in the architecture of the mind. The issue where the two camps disagree concerns representational nativism.

In chapter V I try to explain, first of all, how exactly representational nativism should be explained in more detail, and in what sense we can talk about innate representations that are genetically inherited evolutionary adaptations. It turns out that the best explanation of innate mental representations is offered by what I call Neural Nativism.

Neural Nativism defines mental representations as patterns of cortical activity, which depend on specific patterns of synaptic connectivity. 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 and in specific locations in the brain. According to NN, those pre-specified neural structures are determined to represent specific objects: it has been inherited by the individual and evolved in the species because of its adaptive value: that activation of a given neural pattern constitutes thinking a thought containing a specific thought-part, and so, that it constitutes the occurrence of a given concept. Whenever a given neuronal structure fires up, the organism entertains a given representation.

If we applied Neural Nativism to Fodor's views, we would explain Fodor's mad-dog concept nativism as the position according to which it is a genetically heritable trait and an evolutionary adaptation for the human species 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.

This seems to imply that the way to interpret Fodor's view is as a type-type identity theory according to which every type of mental entity is identical with some type of neural entity. This, however, is not plausible scientifically.

To say that our genes 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 be excited to arouse a given concept. 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 will end up being represented, and the pre-specified neural pattern 'waiting' in the brain, a neural pattern which, when activated, will represent 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.

At this point of its development science is not able to determine whether the view that Neural Nativism promotes is true. All that scientists can determine these days is which parts of the brain are (the most) active during various cognitive tasks. They have no way to say, however, what exactly happens in the brain when I think 'CAT.' So NN is not as of yet supported by science.

Moreover, identity between individual concepts and specific neural patterns in specific locations of the brain would most likely *not* be an evolutionary adaptation. Or at least, it is hard to imagine what evolutionary advantage it would be to have pre-specified neural patterns for each mental representation that the organism could entertain.

What science does tell us is that the same outcome, and so, in particular, the thinking of a certain thought part, 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. It seems to go against scientific evidence to suggest that there have to be exactly the same neural patterns in the same parts of the brain that would correspond to the same concepts in different people.

Concepts as thought-parts are, therefore, in all probability not innate. The situation, however, is not entirely hopeless for Fodor. When we look closer at Fodor's account of concepts, and at his views on the architecture of the mind, we realize that he does not need to keep the scientifically implausible view described above.

Instead, Fodor could 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 don't have to be born with pre-specified symbols of the brain-code. It's enough that we have an innate capacity to 'hire' a neural pattern in response to a given kind of stimulus. We are 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).

This approach would make Fodor's views very similar to those of Aquinas. Both Fodor and Aquinas would agree that a plausible account of cognition cannot be entirely

empiricist. However, the innate elements in cognition that it needs to posit are not innate concepts, but rather innate cognitive mechanisms.

PREVIEW

CHAPTER I

TRADITIONAL INTERPRETATION OF AQUINAS AS AN EMPIRICIST AND ITS FAILURE

1.1 The Aristotelian-Thomistic account of cognition

In part 1.1 of the first chapter I present a summarized version of Aquinas's account on cognition. As we shall see, this account of cognition gives rise to the following four problems: the *Transduction* problem, the *Concepts as Essences* problem, the *Being as the First Intelligible* problem, and the *D/d* problem.

1.1.1 Preliminaries

According to the Aristotelian-Thomistic tradition, the only objects that we encounter in the world are *particulars*, such as an individual pebble, John, an individual person, or Yogi, an individual dog. Any such individual belonging to a natural kind is an instance