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

Thomas Aquinas

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

Aristotle

DA      De Anima
Metaph  Metaphysics
PA      Posterior Analytics
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 Being.\(^4\) ("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.
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.
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
of a substance that is, it is a composite of substantial form and matter\(^5\). Substantial form is what determines the essence (quiddity or nature) of an object, that is, the set of necessary properties responsible for the object's belonging to a given natural kind; it configures matter to be a specific kind of substance, a pebble, a dog, or a man. Matter is the principle of individuation.

In addition to substantial forms, there are also accidental or incidental forms, which determine different properties of substances. Accidents do not exist on their own, but only in substances; they are modifications or attributes of things that can exist on their own, i.e. of things that are already composites of matter and form.

All beings in the natural order form a hierarchy. To what level in the hierarchy a thing belongs is determined by the thing's form. Thus, at the bottom of the hierarchy there are inanimate objects. This particular stone is a piece of granite because its inanimate form organizes its matter (quartz, feldspars and various minerals\(^6\)) into a specific structure and shape. Animate things occupy a higher level in the natural order. Their forms, often called souls, differ from inanimate forms in that they not only organize matter into a specific shape and physical structure but are moreover principles of activity: the soul of any given organic structure determines its powers and capacities for various kinds of behavior. Thus plants, the most primitive among animate beings, are endowed with vegetative souls providing them with powers of nutrition, growth and reproduction (cf. *DA* II 413a19-414a3). Non-rational animals, occupying the next

---

\(^5\) There are, for Aquinas, two main kinds of matter: prime matter and secondary matter. Secondary matter is what science deals with; it refers to different kinds of 'stuff,' or to different kinds of chemical substances that things in our world are made of. Prime matter is pure potentiality.

\(^6\) The elements themselves belong to the very bottom of the hierarchy of things in the natural order.
grade in the hierarchy of beings, possess sensitive souls, which endow them with powers of locomotion, appetite (desire) and perception. The highest place in the hierarchy belongs to human beings who possess rational souls\(^7\) to which Aquinas ascribes the specific powers of will and intellect.

According to the Aristotelian-Thomistic doctrine, humans are a part of the natural order. They share several capacities with non-rational animals. In particular, they share all the powers that are characteristic of the sensitive soul, that is, the powers of locomotion, appetite and perception. At the same time, Aquinas ascribes to humans a special status among all other beings. This is because he believes that the human mind (or reason) is a cognitive power of a different order than cognitive powers possessed by other animals: the human mind for Aquinas is immaterial. A common interpretation of Aquinas's view that the mind is immaterial is that it does not reside in the brain, and that it does not require any physical organ. If this was the correct interpretation of Aquinas, his account of cognition would look quite implausible from the point of view of the contemporary reader. The intellect's immateriality, however, can also be explained as consisting in the fact that the set of things that it can cognize is not determined by any physical features of those objects. The intellect is then contrasted with the senses. Our sense of sight is limited: our eyes can only respond to wavelengths from about 380 to 750 nm. The intellect is not limited in this way: we can cognize anything that is intelligible, no matter what its physical features are.

\(^7\) Each animate being has only one soul; a higher-level organism, in virtue of its higher-level soul, also possesses all lower level capacities. Thus, rational souls not only endow human beings with rational capacities, but also are responsible for their vegetative and sensitive powers.
1.1.2 Stages of cognition

'Cognition,' on Aquinas's account, is a term which applies, properly speaking, to any process in which cognitive faculties of the subject register and process information about the external world. Aquinas distinguishes 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.

(i) Sensation: The External Senses

Human (as well as animal) cognitive contact with the objects in the external world originates with the exercise of the five senses: sight, hearing, taste, smell and touch.\(^8\) The senses interact with external objects in a passive way: a stimulus, a perceptible object, effects a modification in a given sense organ. The object, that is, a composite of form and matter, becomes actually known when a sensible species, that is, the object’s sensible form, is received by the senses of the knower.\(^9\) For example, John becomes

---

\(^8\) I do not mention the very first stage of cognition in which the material object affects the medium between that object and the external senses.

\(^9\) The concept of form, together with the idea that the process of cognition should be understood as a process in which the form of the object cognized is received in the cognizing subject, are crucial for Aquinas's account of cognition. This theory, on its own, appears just unintelligible. It becomes more plausible, however, once we realize that it is "merely a shorthand for some more complicated story" (Hamlyn, 1994, p. 140). A detailed analysis of the account of cognition as acquisition of form will be presented mostly in part 1.2 below.
visually acquainted with an apple that is in front of him when his sense of vision receives (or is modified by) the visible form of the apple.

(ii) Three Kinds of Objects of Sensation

According to the Aristotelian-Thomistic tradition, a cognitive faculty is defined by its objects (cf. DA 415a20, 418a7ff; 1139a8-11; ST I 77, 3). Aquinas find it necessary to distinguish three kinds of objects of the senses (cf. DA II 6 and ST I 17, 2). With respect to any of the senses (or any of the cognitive faculties) we can talk, first of all, of their proper or special objects (proper sensibles). The proper object, as we read in Aristotle, is "what is perceptible by a single sense," and that "which cannot be perceived by any other sense than that one" (DA 418a12). Proper objects of sight are colors, of hearing—sounds, and of taste—flavor. Second, certain properties of things can be perceived by more than one sense. Aristotle calls them common sensibles and lists five of them: movement, rest, number, figure and magnitude (DA 418a18). The third kind of object of perception, incidental sensible, is something that is not perceived directly by the external senses. Aquinas believes that properly speaking we do not see a cat (although it would be correct to say that we perceive it). Vision as such is not sufficient for seeing cats; in itself, vision is a more limited capacity: the capacity to receive color impressions and to distinguish colors. On the other hand,

---

10 Aquinas uses the term 'sensible' as a noun. It means "that which the sense takes in, as distinct from what is cognized by the sense" (Stump, 2003, p. 247). That is to say, when the sense object cognized by the external senses is, for instance, a coffee cup, the sensible taken in by the sense of vision is the cup's color.
Q1. We speak of an incidental object of sense where e.g. the white object which we see is the son of Diares; here because being the son of Diares is incidental to the white which is perceived, we speak of the son of Diares as being incidentally perceived. That is why it in no way as such affects the senses. (DA 418a20-24)

Indeed, even though we might say that we see a man, we see him only in virtue of seeing colors: colors, and not persons, are what make an impression on the eye. This is because the quality of being a man does not, "in its own right, bring about any physical alterations." This quality "may be manifested through sensible qualities," but it is not a sense object "except by accident" (Pasnau, 2002, p. 182).

(iii) Internal Senses: The Common Sense

The next thing that Aquinas, following Aristotle, accounts for is the unified character of experience. Imagine that my dog Yogi is actually seeing something gray, she can hear meowing, and she perceives a middle-sized moving thing. If each of her senses were completely independent from others, the dog would not be able to connect all of the impressions it receives as stemming from one and the same object. To solve this problem Aquinas postulates an integrative capacity called common sense, or sensus communis.

The common sense is not another sense over and above the five external senses. Rather, as Aquinas explains it, it is "the common root and principle of the external senses (…), their common terminus" (ST I 78, 4, ad1-3). The common sense receives sensible forms which are the output of the external senses and unifies them in order to enable an organism to recognize different sensory impressions as originating from one and the same object. The common sense organizes the incoming stimuli into a phantasm, a
unified sensory experience, a unified percept. Another function of the common sense is to make the cognizing subject able to differentiate between the objects of each particular sense (InDA II 13, 390), that is, to make comparisons and distinguish between the objects of different senses, e.g., to discern the white from the sweet.

(iv) Perception: Other Internal Senses

According to Aquinas, in addition to the five external senses and the common sense, cognition requires the functioning of other internal senses. Aquinas, following Avicenna, lists four internal senses: the common sense, phantasia or the imaginative power, the estimative power, and memory (ST I 78, 4). One feature that the internal senses are supposed to account for is the capacity to "apprehend a thing not only when that which is sensible is present, but also when it is absent" (ST I 78, 4). Both sensible species and phantasms (i.e., vehicles of cognition produced and operated on by both external and internal senses) must be stored somewhere for the animal to be able to recall them, to recognize elements of past experience, to learn and set goals. Aquinas holds that it is the internal sense of imagination (that he also calls phantasia) that retains sensible species when the sensible objects are no longer present. The internal senses of the vis memorativa (sense memory) serves as a storehouse for phantasms produced by higher level internal senses.

---

11 For a more thorough explanation of Aquinas's views on the internal senses see below, 2.3.
12 Aristotle talks about imagination and memory in addition to the common sense (although he does not refer to them as 'inner' or 'internal' senses).
In addition, the cognizing subjects need a faculty that will enable them to recognize various things that the external senses by themselves cannot perceive. Specifically, they need the capacity to distinguish what is dangerous for them from what is beneficial, what should be pursued from what should be avoided.\footnote{There is a debate among scholars as to whether Aquinas (and Aristotle) ascribes to the internal senses yet another role, that of what today we call consciousness. It may seem, indeed, that the functioning of the external senses together with the common sense, memory and imagination, put in terms of reception of sensible species, i.e. registering of sensory impressions, is still not sufficient for genuine cognition, and that it should be completed by another capacity on the part of the subject, that of being aware of, or having a conscious access to the data processed by the senses. Thus, Eleonore Stump suggests that for Aquinas it is the functioning of \textit{phantasia} that provides the subject with awareness of its mental states (e.g., Stump, 2003, p. 267). Robert Pasnau (Pasnau, 2002, pp. 196–9) and Charles Kahn (Kahn, 1979, pp. 10–11), on the other hand, think that it would be more plausible to ascribe this role to the common sense (the former refers to various quotes from the \textit{Summa Theologiae}, the latter to Aristotle's views in \textit{Parva Naturalia}). At the same time, it may seem that the problem of consciousness simply does not concern either for Aristotle or for Aquinas. The two philosophers do not speak the language of dualism, i.e., they do not distinguish between the material body and the realm of the mental. It would not occur to them to think about sensations, or mental states in general, that are \textit{unconscious}. Thus, perhaps they'd hold that there is simply no special faculty responsible for consciousness, nor a special place in which it occurs. Instead, it might be present everywhere in the process of cognition. A newborn baby or a person waking up from a coma could be aware of some vague and indefinite stimuli—this would be the case of consciousness at the level of the external senses. The internal senses, on the other hand, could be said to bring with themselves a more genuine conscious awareness of things and their properties. In any case, the problem of consciousness does not really matter for the purposes of this dissertation.} As Aquinas puts it,

Q2. It is necessary for an animal to seek or flee from things not only because they are or are not agreeable to the senses, but also for the sake of some further benefits and uses, or harms. Thus the sheep flees when it sees the wolf, not because its color or shape is unattractive, but as if the wolf were harmful to the sheep's nature [\textit{non propter indecentiam coloris vel figurae, sed quasi inimicum naturae}]. (ST I 78, 4)

This function, on Aquinas's account, is performed in non-rational animals by the internal sense faculty of the estimative power. A counterpart of this internal sense in humans has the name of the \textit{vis cogitativa}. The cogitative power turns out to be crucial for Aquinas's account of concept acquisition (see below 2.4.3 and especially 3.3).
(v) Thought

The final stage of cognition is exclusive to beings endowed with the rational soul: it takes place at the level of the intellect. According to Aquinas, the intellect can operate in three different ways. The first operation of the intellect consists in forming an abstract idea, or a concept, of a given object of cognition. At the second level, association of concepts takes place. In the last stage the intellect engages in discursive reasoning.

(Before we move on, it is important to notice already at this point that Aquinas actually uses two terms, *intelligible species* and *mental word*, to refer to vehicles of cognition at the level of the intellect: both terms are sometimes used interchangeably with the term 'concept.' We'll see below, especially in part 3.1, that a clear explanation of the difference between intelligible species and mental words is crucial if we want to ascribe to Aquinas a plausible theory of concept acquisition. I argue that it is really the intelligible species that should be understood as concepts, at least if we take concepts to be thought-parts, which is a view held by Fodor, and also a view that, as we shall see, can plausibly be ascribed to Aquinas; see 3.3.)

In order to cognize a thing the intellect must somehow actually become that thing. As Aquinas puts it following pre-Socratics and Aristotle, "a thing is known as far as its form is in the knower" (*ST* I 75, 5), and a bit further, "the likeness of the thing understood is the form of the intellect" (*ST* I 85, 2). Again, as it was with the senses, so

---

14 Cf. 1.2.3 (iv) for a discussion about the process of abstraction, and 3.1 for a detailed account of what concepts are for Aquinas.
also here: what the intellect acquires is the form of an object. As Aquinas puts it, "the soul is not simply identical with the things it knows; for not stone itself, but its formal likeness exists in the soul" (InDA III 13, 789).

The intellect differs from the senses first of all in that it is a single faculty with two distinct powers: the active and the passive power. Things in the external world are not actually intelligible. It is the role of the agent intellect in its first operation to act on the phantasm produced in the internal senses and to highlight those features in the phantasm that are capable of being captured in an intelligible species, or a concept, while ignoring those features of the thing that are irrelevant. This operation is called abstraction. The intellect abstracts intelligible species from the material conditions (cf. ST I 79, 3), and then on this basis it produces mental words. These are then stored in the passive (or possible) intellect.

The second operation of the intellect is called compositio et divisio, that is, 'putting together and taking apart' (cf. e.g., DA 430a26-b6 and InDA III 11, 746). This is where the association of concepts takes place and where the intellect formulates positive or negative judgments about things cognized. This is also where truth and falsity enter the stage because here "thoughts are considered in their relation to reality" (InDI I 3, 26). Finally, the last operation of the intellect is discursive reasoning in which the intellect makes inferences and broadens the scope of knowledge.16

15 For Aquinas "to think (cogitare) is, as it were, to shake together (coagitare), that is, to separate and compare one thing with another" (DV 24, 1).
16 The account of the last two stages of intellectual cognition that I provide is quite sketchy. It is sufficient, however, for the purposes of this dissertation.
The intellect, for Aquinas, differs from the senses also in several other ways. The most important difference for our purposes is that while the senses "have cognition of external accidents only, intellect alone succeeds in reaching a thing's essence" (DV 1, 12). Universal essences, abstracted from their material conditions (cf. ST I 79, 3), are the proper object of the intellect. In addition, Aquinas also says that "that which the intellect first conceives (…) and to which it reduces all its concepts is Being" (Illud autem quod primo intellectus concipit (…), et in quod conceptiones omnes resolvit, est ens; DV 1, 1). Being, for Aquinas, is the first intelligible (ST I 5, 2: ens (…) est primum intelligibile) without which "nothing can be apprehended by the intellect" (In I Sent 8, 1, 3: nihil potest apprehendi ab intellectu).

1.1.3 The mechanism of cognition

(i) Form

A crucial element of the Aristotelian-Thomistic account of cognition is the idea that both human beings and non-rational animals cognize the world through the process of the acquisition of form. At all levels, cognition consists in a reception of an object's form by the cognitive faculties of the cognizer. The reason why this form, in the context of cognition, has the name of species is that it specifies a given content, or a given feature that becomes known to the cognizing subject (cf. Deely, 2007, p. 24). Properly speaking, we should distinguish two different kinds of species in cognition: species impressae which
stand for what stimulates a given cognitive faculty, and species expressae which stand for
the response, or the output of the cognitive faculty (cf. Deely, 2007, p. 27). The species
determine what objects (or features of objects) in the external world the cognitive
faculties will identify.17

The cognitive process, properly speaking, goes as follows. Cognitive contact with
the objects in the external world originates with the exercise of the five external senses:
sight, hearing, taste, smell and touch. At the first level of cognition, the species impressae
of the external senses, that is, some perceptible feature of an object, a stimulus in the
form of a proper or a common sensible, effects a modification in a given sense organ. As
a result, the species expressae of external senses, or sensible species, are formed. At the level
of the internal senses, the species expressae of perception, or 'phantasms,' are formed.
After that, the agent intellect transforms phantasms into species impressae intelligibilis
(Deely, 2007, p. 53), or intelligible species. (Intellectual cognition, for Aquinas, is
impossible without some kind of 'return to phantasms.' More on this below.) And
finally, the possible intellect, stimulated by the intelligible species, responds by creating
the species expressae intellectae, or the mental word.

17 Cf. Deely, 2007, p. 27: Species impressa specifies "what the sense will objectify in its interaction with
the physical surroundings (...); it's the effect of [physical] stimulus here and now active upon a sense organ and
placing it in relation with the here and now active source existing on the same material level as the sense
organ itself (...). The 'form' carried by the stimulus specifies the response of the organ to objectify this rather
than that aspect of the surrounding environment." There is a lot of information coming from the external
world; cognitive forms include only some of it.
On one understanding of the term, a *form* is a pattern, or a code in which all the necessary information about the individual substance (or its accidents) is written.\textsuperscript{18} The form of an object can be understood as "the configurational state in which the matter of that object is arranged," similar to the configurational state of protein that "is preserved in the code of DNA" (Stump, 2003, pp. 249, 252–3). The role of form in general, that is, both in cognition and independently of it, is to "bring a certain identity to things, causing them to be units of *a certain kind*" (O'Callaghan, 2003, p. 238). In reality, the form makes a thing be what it is. In cognition, it determines the object that the knower cognizes. Different forms of the object cognized are represented by different vehicles of cognition in an encoded fashion (on the relationship between cognitive forms and forms of objects of cognition see below, esp. 1.2.).

In the Aristotelian-Thomistic tradition, the form, as Anthony Kenny puts it, is "the real object of all human knowledge (...) This is true of sensory acquaintance and intellectual understanding" (Kenny, 1980, p. 73). Such a claim may not sound plausible to the contemporary reader. After all, we cognize material objects in the extra-mental reality, and not forms. Forms can be said to be the object of human knowledge only in the sense that all cognitive faculties operate on them. Aquinas explicitly rejects the view according to which our cognitive powers "cognize nothing other than their own states—for example, that a sense senses nothing other than the state of its organ," or that the intellect "understands nothing other than its own state—that is, the intelligible species

\textsuperscript{18} Aquinas says that a form can be understood also "in its broader sense as meaning an intelligible character" (*DV* 3, 5).
received in it" (ST I 85, 2).\(^{19}\) Forms, for Aquinas, including sensible species, phantasms, and intelligible species, are likenesses of things in the extra-mental reality;\(^{20}\) they are not what we cognize but rather that *by which* we cognize; they are the *means* of cognition. As Aquinas explains,

Q3. The intelligible species, by which possible intellect becomes actualized, are not intellect's object; for they are related to intellect not as what is intellectively cognized but as that by which intellect cognizes. In the same way, too, the species that is in the sense of sight is not what is seen but that by which the sense of sight sees. What is seen is the color that is in a body. Similarly, what the intellect cognizes is the quiddity that is in things, but not the intelligible species (except insofar as intellect reflects on itself). (*InDA* III 8, 718)\(^{21}\)

(ii) Two Kinds of Change and of Reception of Form

The capacity to receive forms is necessary but not sufficient for cognition. There are, in fact, two different ways of receiving forms, natural and spiritual, corresponding to two kinds of change. Aquinas explains:

Q4. There are two kinds of impressions, one natural and the other spiritual. A *natural impression* occurs inasmuch as the form of what does the impressing is received in the thing impressed upon according to its natural existence, as heat is received in the thing being heated. A *spiritual impression* occurs inasmuch as the form of what does the impressing is received in the thing impressed upon according to its spiritual existence, as the form of color is received in the pupil, which is not by this made colored. Sensory operation requires a spiritual impression through which an intention of the sensible form is produced in the organ of sense. Otherwise, if an entirely natural impression were to suffice for sensation, then all natural bodies would sense when they are altered. (*ST* I 78, 3)

---

\(^{19}\) For a discussion on both a representationalist and a direct realist reading of Aquinas see Panaccio, 2001.

\(^{20}\) Sensible species, for instance, is defined as the impression made on the external senses by a material object (cf. *ST* I 78, 3).

\(^{21}\) The intelligible species or concepts can become the object of the intellect's knowledge because the intellect is a reflexive power: it can reflect on its own acts and acquire a second order knowledge about them.
The imposition of form on matter, that is, the natural or material reception of form, is what explains the phenomena of change, generation and corruption. Acquisition or loss of the substantial form account for, respectively, coming into being, and destruction of a thing. When a thing gains or loses an accidental form, on the other hand, it merely undergoes change, and still remains the same thing.

In cognition, the form of an object is received in the senses and in the intellect of the knower in a *spiritual* or *immaterial* way (in this context we can also talk of intentional, or de-materialized reception of form). In reality, the form is what makes a thing a cat, or what makes a cat black. In cognition, the knower receives the form 'cat,' or the form 'black,' without herself becoming a black cat.

According to Aquinas, in the external senses, which use corporeal organs, the form of the object cognized can be received either solely in the spiritual way or in both spiritual and material ways at the same time. The latter happens, for instance, during the reception of sound. When I listen to Mozart's *Requiem*, my hearing the music is an instance of immaterial reception of form: my ears do not acquire the feature of 'sounding like the *Requiem*.' However, my hearing the music is possible only because a natural reception of form (i.e., a material modification of the sense organ) is happening at the same time, that is, vibrations of the air cause vibration of the inner ear. Nonetheless, it is clearly the spiritual or intentional reception of form that is crucial for cognition.22

---

22 'Natural' or material reception of form is the only way in which forms can be received by both inanimate beings and those solely endowed with the vegetative soul. At these levels in the hierarchy of beings there could be no sensation, and no cognition.
(iii) Different Levels of Dematerialization of Form

Different stages of cognition involve different levels of dematerialization of form. In *De Veritate* Aquinas explains:

Q5. in the sensible thing the species has an extremely material existence, but, in the understanding, a very highly spiritual existence. Hence, it has to pass to this spirituality through certain intermediate levels, inasmuch as it has a more spiritual existence in sense than in the sensible thing, and a still more spiritual existence in imagination than in sense, and so on as it goes higher. (*DV* 19, 1)

The physical external object itself has its individual, enmattered form (as the form of *this cat*). Then, in the external senses, the form of the object is still a particular form (as when I see the color of fur or hear the meowing of *this cat*) and exists under material conditions—senses are bodily organs stimulated by physical input (the transmission of photons in the case of sight or of sound waves in the case of hearing). In the final stage of cognition, Aquinas believes, the concept of 'cat' has a different, universal existence; it is completely independent from any matter. The form 'catness,' that is, the *quiddity* of the object *cat* as it exists in the mind "includes only what defines the species of a thing" (*ST* I 3, 3). (An individual cat, then, possesses something which his feline nature does not.) It is not entirely clear what makes it the case that it is still the same form that first configures matter into an individual cat, and then is processed by cognitive powers in different stages of cognition. (This issue will be discussed below, especially in part 2.2. For now we could perhaps say that this is due to the fact that at every level the form

---

23 Different stages of cognition that I also call different levels of dematerialization of form, are often referred to as different stages of *abstraction*. I prefer, however, to keep the term 'abstraction' for the activity of the agent intellect which, "by abstracting the species from material conditions" (*ST* I 79, 3), makes the natures of things actually intelligible. For more on abstraction, see esp. 1.2.3 (vi).
includes some specific encoded information referring to the original material object that activated the process of cognition in the first place. Some pieces of information about the object appear at every level of cognition. This is what allows for cognition of one given object throughout the whole cognitive process.)

(iv) Between the Senses and the Intellect: Four Problems

In the *Summa Contra Gentiles* Aquinas says:

Q6. Our intellect's knowledge, according to the mode of the present life, originates from the senses: so that things which are not objects of sense cannot be comprehended by the human intellect, except in so far as knowledge of them is gathered from sensibles. (*SCG* I 3, 3)

Aquinas holds that our intellect, which operates on intelligible species, i.e., on universal forms, somehow takes up its objects of cognition from the senses. That for Aquinas sensation is a necessary first stage of cognition is quite obvious and unquestionable. The problems begin, however, once we try to figure out how it is possible that our universal concepts take their origin in sensation. For Aquinas, intelligible species, or concepts, are universal and immaterial forms which represent the essences of things cognized (cf., e.g., *ST* I 84, 1, ad1). Sensation, on the other hand, concerns only what's particular and material; its vehicles of cognition, i.e., sensible species, are cognitive forms that represent accidental features of things. How can they, then, give rise to concepts?

---

24 See more on this distinction below, in 1.2.3 and 2.1. Also, cf. 1.1.3 where I talk more in detail about the role of forms in cognition.
There are, in fact, four serious problems threatening Aquinas's account of cognition. These problems indicate the main concerns of this dissertation (two additional problems are introduced in chapter III).

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, that 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 (ST I 5, 2) and that "without Being nothing can be apprehended by the intellect" (In I Sent 8, 1, 3). These claims, first of all, sound very mysterious. Also, given that, for Aquinas, essences in the mind are always universal (in the sense that numerically distinct things can have the same essence), and that 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. If essences are the only kind of thing that the intellect can operate on, then it doesn't seem plausible to say that Being is the first intelligible.

P3. What constitutes the most challenging problem in Aquinas's account of cognition is the issue of concept acquisition. It is not clear how sensible species are supposed to

---

25 This is why it is possible to call both an earthworm and God—a Being.
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.* The crucial question that we'll consider is, therefore, whether Aquinas provides a plausible account of the passage from the sensible to the intelligible, from the non-conceptual to the conceptual.

P4. Finally, Aquinas's account of cognition also faces what Jerry Fodor calls the doorknob/DOORKNOB problem, or the d/D problem (cf. 4.3.5 (i)). As Fodor notices, the relationship between sensory experiences and concepts produced in our mind appears to be entirely random: there seems to be no way to explain why having *those* experiences leads to possession of *this* concept. Usually we think, Fodor explains, that "what leads to acquiring a concept is *typically having the right kinds of experiences*" (Fodor, 1998, p. 127). What we do not know, however, is why it is, "so often experiences of doorknobs, and so rarely experiences with whipped cream or giraffes, that leads one" to acquire the concept DOORKNOB.

In the remaining part of this chapter we are going to look at the common interpretation of Aquinas's views on cognition. We'll see that even though this interpretation doesn't have to face the transduction problem, it turns out to be implausible for other reasons. A possible solution to the transduction problem will be indicated in the second chapter and further developed in the third. Solutions to the *Concepts as Essences* problem and the *Being as the First Intelligible* problem will be offered.

---

26 See King, 1994.
in the third chapter. Finally, a suggestion on how the d/D problem is not really a problem for Aquinas will be presented in chapter V.

1.2 Failure of the traditional interpretation of Aquinas as an empiricist

As we said, cognition for Aquinas is a matter of acquisition or assimilation of the form of the object cognized by the cognizing subject. It is not clear, however (nor is there an agreement with respect to), how this acquisition or assimilation of form is to be understood. In addition, it is not clear what the form in cognition actually is or what its relationship to the form of the object cognized is. These issues will be discussed in this part of chapter I.

1.2.1 The Form Transmission account — A common (mis)interpretation

Aquinas holds that that

Q7. knowledge is an assimilation of a knower to a thing known (DV 2, 2, obj. 3).

Q8. intelligent beings are distinguished from non-intelligent beings in that the latter possess only their own form; whereas the intelligent being is naturally adapted to have also the form of some other thing; for the idea of the thing known is in the knower (ST I 14, 1).

Q9. a thing is known in as far as its form is in the knower (ST I 75, 5).
Q10. [when we cognize a stone] the form of a stone (…), as to its proper formal idea, is in the intellectual soul (ST I 75, 5).  

Aquinas also says that "all our knowledge originates from sense" (ST I 1, 9) and, more specifically, that "every species through which our soul understands is abstracted from sensible things" (DV 10, 8, otc9). These quotes suggest that Aquinas is an empiricist with respect to cognition. They are commonly interpreted as expressing the idea that the very form, which informs the external object and determines it to be the kind of thing it is, (1) first—as sensible species—affects the external senses, then (2) appears as a phantasm in the internal senses, and eventually (3), as intelligible species, is impressed onto the possible intellect. This, supposedly, explains how sensible species 'give rise' to concepts: in some sense, they are one and the same thing, only in different cognitive faculties. In Eleonore Stump's words,  

Q11. the form which is the sensible species is preserved in the phantasm, and the agent intellect abstracts that form from the phantasm in order to make possible intellectual functioning. (Stump, 2003, p. 257)  

Etienne Gilson explains that the role of the agent intellect is "to furnish the possible intellect with the intelligible and universal which lay in the sensible" (Gilson, 1961, p. 218; my emphasis). The same point is made in an even more explicit way by Robert M. Adams when he explains that  

27 See DA III 3, 429a14-15: "The thinking part of the soul (…) must be potentially identical in character with its object without being the object." Cf. also Aristotle, Metaph XII 7, 1072b20-25: "Thought thinks on itself because it shares the nature of the object of thought; for it becomes an object of thought in coming into contact with and thinking its objects, so that thought and object of thought are the same." Cf. also ST I 84, 1 and ST I 85, 2.
Q12. One and the same form, originally present in the object, is present also in the medium, the sense organ, and the mind. There is something (the sensible form) which literally comes into the mind from the object. (…) Intelligible forms (such as those of body, apple, dog, and being) also come into the mind from outside, in the sensible forms. (Adams, 1975, p. 73)\(^\text{28}\)

The passages quoted from Aquinas scholars suggest that it is in some way the same form that is processed by different faculties in both sensory and intellectual stages of cognition. This agrees well with the empiricist slogan according to which there is nothing "in the intellect that was not previously in the senses" (DV 2, 3, obj. 19).

Now, if cognition consists in the assimilation or reception of the form of the external object into the cognizer's mind, as we read in the passages from Aquinas's own texts, this suggests that not only do we deal with sameness of form in the different cognitive faculties, but also that the form of the object cognized must be in some sense the same as the form in the cognizer's mind. This seems to suggest an even stronger empiricist claim to the effect that there is nothing in the intellect that was not previously in the object itself. It is far from obvious, however, what meaning should be ascribed to the phrase 'the same form' in this context. In what way could two forms, one in an extra-mental object, and the other in the cognitive faculties of the cognizing subject, be the same? What meaning does 'the same' have anyway? For instance, what does a person mean when she says that the coin in her pocket and the coin that she's thinking about are the same? On the most radical, or strongest, understanding of the term, the coins are the same if we deal with numerical identity. In that case, you couldn't have in your pocket two coins that are the same, but really only one coin. Another possibility is that

\(^{28}\) Cf. also Haldane, 1983, p. 235.
two coins are the same if they are type-identical, e.g., if both are pennies. In that case, however, we could still ask how similar two coins have to be for us to call them the same. Are they still 'the same' if one is a penny made of copper and the other of gold? Similar questions can be asked with respect to the issue of identity, sameness or similarity of forms.

It is not clear what Aquinas himself means when he says that the form of an object known is in the knower (ST I 75, 5). He explains that

Q13. an external thing understood by us does not exist in our intellect according to its own nature; rather, it is necessary that its species be in our intellect, and through this species the intellect comes to be in act. Once in act through this species as through its own form, the intellect knows the thing itself. This is not to be understood in the sense that the act itself of understanding is an action proceeding to the thing understood, as heating proceeds to the heated thing. Understanding remains in the one understanding, but it is related to the thing understood because the abovementioned species, which is a principle of intellectual operation as a form, is the likeness of the thing understood. (SCG I 53, 2)

Aquinas explicitly calls the form in the mind a 'likeness' of the form informing the object. Nonetheless, the way he talks about the species of the thing being in the intellect leaves room for different interpretations of how this should be understood. What these interpretations seem to have in common is that they all explain cognition in terms of some kind of a passage, or transmission, of form from the object through the senses, and to the intellect. This is what is called the form-transmission account of cognition (from now on, the FT model or FT interpretation).²⁹

I will discuss three such interpretations — three versions of the interpretation of Aquinas's account of cognition in terms of acquisition of form. The interpretations differ

²⁹ Cf. Adams, 1975, p. 73, and Jacobs and Zeis, 1997, p. 541 where they talk about the transmission of form.
in their understanding of the exact nature of the relationship between the form informing the object and the form in the mind, and in their conclusions with respect to empiricism/nativism debate. Two of the interpretations that are the most common, that is, numerical identity theory and formal sameness theory, I am going to reject. The third one—that I am going to call the *Form Trans-Formation Theory* (the FTF) and also the *Similarity Theory*\(^{30}\)—will emphasize the need to posit certain innate elements in Aquinas’s account of cognition.

(A) Numerical Identity version of FT (it is called ‘Identity Theory’ in Brower, 2008; in King, 2005 it corresponds to R1, and specifically to what he calls the simple version of Conformality Theory);

(B) Formal Sameness version of FT (called ‘formal-sameness theory’ in Brower, 2008; cf. also R2 in King, 2005);

(C) Similarity interpretation— the FTF account of cognition.

(i) Numerical Identity of Forms

(A) On the Numerical Identity version of the FT, the form that informs the object in the external world, and the form that is the concept in the cognizer's mind are the same in the sense that they are *numerically identical*. This means that there is *numerically one* and the same form as a concept in the intellect and as the form of the material thing in the world. On this view, "a mental representation represents an object just in case it has the

---

\(^{30}\) In a way, the Form Trans-Formation account (the FTF) can be considered to be another version of the FT interpretation: they share the idea that cognition consists in some kind of passage of the form. If not indicated otherwise, however, I reserve the name ‘the FT model’ for the two interpretations of Aquinas that I reject, as opposed to the Form Trans-Formation account (the FTF) that I propose.
same form as the object" (King, 2005, p. 82), or perhaps just in case it is the same form as the form of the object.

It should be immediately obvious that this version of the FT account that is based on the most literal understanding of sameness of forms, faces serious problems. We said above that form is that which makes something to be the kind of thing it is. If that's the case then it is not clear why the mind itself would not become a cat when it receives the form 'cat.' In addition, if the form is exactly the same throughout the whole process of cognition, then there seems to be no reason for Aquinas to talk about three distinct kinds of cognitive faculties or three different levels of cognition. Finally, there is a problem with the very idea of identity of forms involved in the process of cognition, given the fact that identity is a symmetrical relation. Forms in the cognitive faculties are supposed to represent substantial and accidental forms of the external objects. But if what we deal with here is identity of forms, then it seems that the substantial and accidental forms of the object should also be said to represent the forms in the mind. If cognition occurs in virtue of identity of forms, then cognition itself should be symmetrical, and this seems counterintuitive. (Similar problems also threaten the Formal Sameness version of the FT. They are discussed in more detail below, in 1.2.3.)

What is the most important is that the Numerical Identity interpretation does not seem to agree with Aquinas's own understanding\(^{31}\) of the relationship between the form of the object cognized and the form in the cognizer's mind. Aquinas does talk about the

\(^{31}\) Brower (2008) ascribes this view mostly to certain textbook accounts of Aquinas (Crane, 1998 and Cummins, 1989), but finds it also, more surprisingly, in Geach (1961) and Haldane (1998).
form in the cognizer's mind being similar to (or being a likeness of) the form of the object
cognized. He also believes that things are similar if they share forms. He is aware,
however, that this can happen in a number of ways. As he explains,

Q14. since likeness is based upon agreement or communication in form, it varies according to
the many modes of communication in form. (ST I 4, 3) [cum similitudo attendatur secundum
convenientiam vel communicationem in forma, multiplex est similitudo, secundum multos modos
communicandi in forma.]

The passage continues:

Some things are said to be like, which communicate in the same form according to the same
formality, and according to the same mode; and these are said to be not merely like, but
equal in their likeness; as two things equally white are said to be alike in whiteness; and this
is the most perfect likeness. In another way, we speak of things as alike which communicate
in form according to the same formality, though not according to the same measure, but
according to more or less, as something less white is said to be like another thing more white;
and this is imperfect likeness. (ST I 4, 3)32

There are, therefore, for Aquinas, different kinds of relationships between forms that
could still be called similarity, likeness, or sameness. If we encounter, for instance, 'two
things equally white,' we would consider them 'alike in their whiteness.' In this case, it
would be correct to say of two such things that they possess 'the same' form white.

32 The rest of the passage goes as follows: "In a third way some things are said to be alike which
communicate in the same form, but not according to the same formality; as we see in non-univocal agents.
For since every agent reproduces itself so far as it is an agent, and everything acts according to the manner
of its form, the effect must in some way resemble the form of the agent. If therefore the agent is contained in
the same species as its effect, there will be a likeness in form between that which makes and that which is
made, according to the same formality of the species; as man reproduces man. [This is what happens when
one concept, or one mental representation, leads to another.] If, however, the agent and its effect are not
contained in the same species, there will be a likeness, but not according to the formality of the same species;
as things generated by the sun's heat may be in some sort spoken of as like the sun, not as though they
received the form of the sun in its specific likeness, but in its generic likeness. [This would be the case of an
object in the external world causing a mental representation in the cognizer's mind.] Therefore if there is an
agent not contained in any 'genus,' its effect will still more distantly reproduce the form of the agent, not,
that is, so as to participate in the likeness of the agent's form according to the same specific or generic
formality, but only according to some sort of analogy; as existence is common to all. In this way all created
things, so far as they are beings, are like God as the first and universal principle of all being." (ST I 4, 3)
Notice, however, that even here where we deal with what Aquinas calls 'the most perfect likeness' of forms, the 'sameness' of the two forms is not numerical—we have, after all, two things. Rather, each has the same form 'white' only in the sense of type-identity (cf. Stump, 1998, p. 289, and pp. 304–5). Even the 'most perfect likeness,' therefore, does not refer to numerical identity. Let us see, then, if it is more plausible that what Aquinas has in mind when he talks about the similarity of forms is formal sameness between two things.

(ii) Formal Sameness

(B) On the Formal Sameness version of the FT, the form of the object cognized, and the form in the cognitive faculties are only formally the same; they are not numerically one; rather, they could be seen as type-identical. This version of the FT model would explain cognition as just scanning and copying: the intellect takes the form existing as a phantasm and copies it for its own use, while at the same time stripping it off its material conditions. This activity does not require that the intellect do any sort of 'creative' work. The intellect does not transform in any way the content of the form that it finds in the phantasm.

---

33 For Aquinas, forms in reality are always individual, they are individualized by the matter that they inform. As he puts it when he talks about the form 'humanity,' "humanity is something in reality, but there it is not universal, for no humanity outside the soul is common to many" (In I Sent 19, 5, 1). Numerical identity of forms involved in cognition would require some kind of realism about universals, and this is a view that Aquinas clearly rejects. Cf. Brower, 2008, p. 19.
If we use Aquinas's vocabulary, we'd say that we deal with formal sameness in cognition when the form informing the object of cognition and the form as the corresponding concept are identical in their intelligible character. Aquinas explains:

Q15. Although in the mind there are only immaterial forms, these can be likenesses of material things. For it is not necessary that the likeness and that of which it is the likeness have the same manner of existing, but only that they agree in intelligible character, just as the form of man in a golden statue need not have the same kind of existence as the form of man in bones and flesh. (DV 10, 4, ad4)34

In order to clarify the notion of agreement of forms in their intelligible character we could think of the relationship between a city and a map of the city. It makes sense to say that the map successfully represents the city if they share the same form. The 'sharing of the same form' that we talk about in this context, however, is rather metaphorical. The city is three-dimensional, and the map is two-dimensional. There are different elements on the map (dots, squares, lines, etc.) and different elements that make up the city (buildings, streets and cars). Also, the elements on the map are related to each other in a different way than the landmarks of the city are related to each other. At the same time, we can talk of the map representing, successfully or not, the city, and we can build a city according to a map. This is because there is something that they have in common, there is some kind of correspondence of form between the city and the map; the configuration of the elements on the map corresponds to the configuration of the landmarks of the city (cf. Sellars, 1963, pp. 41ff)35.

35 Also, consider a similar explanation offered by King: "Imagine a monochrome sketch of a sheep. The sheep isn't literally colored a shade of gray, but that gray shade rather than another corresponds to the
The correspondence between forms does not hold between the map itself as a physical object, a piece of paper with some ink on it, and the buildings and streets of the city. Similarly, it is not the concept itself as a mental item that we’d call formally the same as the object's form. A concept as a mental object is a quality, and a quality can only be the same as another quality (cf. Brower, 2008). If we want to talk about the relationship of sameness between the two forms, we should perhaps say that it is what the vehicle of cognition represents that is formally the same as the form of the object.

Recall that on one understanding of the term (cf. 1.1.3(i)) a form can be taken to be a code (similar to the genetic code) in which all the necessary information about an individual thing is written, a code which stands for the configurational state in which the matter of an object is arranged. If this understanding of form works, then we could say that on the Formal Sameness version of the FT, the form in the intellect (Fi), the form in the senses (Fs), and the form of the object (Fo), are formally the same in that they all encode the same information about the same configurational state in which the matter of the object is arranged. The role of the intellect, on this account, would be to translate information contained in a phantasm, that is, information that the senses can read, into information that is processed by the part of the brain that is responsible for thinking and higher intellectual cognition. The intellect would only have to change the format of the encoding.

sheep's rich and creamy merino color. Likewise, the sheep isn't a two-dimensional closed plane figure, but is sketched as one according to the 'laws' of perspective. In short, none of the sheep's properties are literally present in the representation; the representation, instead, pictures what it does in virtue of having intrinsic features that correspond to the properties of the external object. The image need not 'look like' its subject at all, as long as the appropriate correspondence holds” (King, 2005, pp. 91–92).
As we shall see below (1.2.3), Formal Sameness version of the FT will have to be rejected because it also faces the problems mentioned above with respect to Numerical Identity version.

(iii) The FTF or the Similarity Version of the FT

(C) The interpretation of cognition that I propose, the Form Trans-Formation interpretation (for a more detailed account, see below, 2.2), can also be called the Similarity account of cognition. On this account, even though there is neither numerical, nor formal identity between the forms involved in the cognitive process, the involvement of forms is still required for cognition (this is why the FTF could still, in principle, be considered a version of FT). Also, there still has to be a link between the form in the mind and the form informing the object in the external world. We call these forms similar just to make sure that we keep Aquinas's terminology: he is after all quite explicit in his use of the term 'similarity' for the relationship between the form of the object cognized and the form in the cognizer's head. But 'similarity' in this context has a really different meaning from what we'd usually take it to have.36 It refers to a special kind of causal connection between forms involved in cognition.

In what follows, I first consider the reasons for which Aquinas scholars hold on to the FT interpretation. Then, I discuss in more detail the problems that make the FT interpretation not acceptable.

36 It’s important to keep in mind that my understanding of the similarity relation between forms is very different from, e.g., Brower’s, in his 2008, pp. 27–28.
My claim is that Aquinas scholars who hold on to the Numerical Identity or to the Formal Sameness versions of the FT interpretation do it not only because this reading seems to be suggested by those texts of Aquinas's where he follows Aristotle. They have, at least, two additional reasons. A weaker reason is that they hope that the FT will help prevent the post-Cartesian worries about metaphysical idealism (MI) and skepticism. A stronger reason is that they hope that the FT actually guarantees metaphysical realism (MR), a position that is extremely important for Aquinas. As it turns out, however, Aquinas himself never felt threatened by skepticism or MI, and he doesn't need the FT to keep his MR.

1.2.2 Reasons behind the FT interpretation

The FT interpretation, if it was defensible, would have a serious advantage in that it faces neither the Transduction Problem (P3), nor the d/D Problem (P4 as described in 1.1.3 (iv)). If the form at all levels of cognition is the same, then there is no problem with the passage from the level of the senses to the level of the intellect. This does not seem to be, however, an explicit reason for its popularity.

The fact that the FT interpretation, in either Numerical Identity or Formal Sameness version, is the most common reading among Aquinas scholars, has something to do, it seems, not only with the scholars' conviction, first of all, that this is what is implied by the terminology that they find in Aquinas's texts, but more importantly with
their desire to present Aquinas as a metaphysical realist. Before I identify and evaluate
the reasoning behind this interpretation of Aquinas, let me first provide brief definitions
of the terms involved as I use them.

Most generally, while realism and idealism are metaphysical positions concerned
with the nature of reality, with what exists, empiricism and nativism are epistemological
positions that deal with the nature of human cognizers and their ways of cognizing.
Metaphysical realism (MR) asserts (and metaphysical idealism rejects or at least
suspects) the claim that there exists a reality, i.e., material objects and their qualities,
prior to and independent of any cognizers. Metaphysical realism is typically
accompanied by an additional epistemic thesis to the effect that we have access to and
actually cognize this kind of external and independent reality. It emphasizes that
objects of human knowledge have an existence that is prior to and independent of
whether they are being cognized by anyone. Metaphysical idealism (MI) typically
includes the epistemic claim that the mind does not have access to any kind of
independent reality but rather it can grasp only its own contents, e.g., only its own sense
perceptions.

37 The latter is quite explicit for instance in Sellars (cf. Sellars, 1963) who seems to believe that the advantage
of the FT is that it actually guarantees MR. (Cf. also Haldane, 1995, p. 21 and Brower, 2008, pp. 25 and 18.) It
should be clear, however, that this, most likely, was not Aquinas’s own concern, but only his scholars.’ I’ll
come back to this issue below.
38 The two claims, i.e., that there exists a reality independent of any cognizers and that we actually have
access to this reality, do not have to go together; one could believe—like Kant—that there exists such an
independent world, but that we don't have access to it in any way.
A cognitive empiricist (CE) in an extreme version of the view claims, roughly, that all our knowledge, beliefs and concepts come from experience. Every person is born with an empty mind, a tabula rasa, and what she gets to know (to think, or to believe) is entirely determined by what she experiences. (A moderate cognitive empiricist would agree that what concepts or beliefs get acquired depends also on innate differences in the ways in which the external senses work in different people, on differences in their innate intellectual capacities, and so forth.) A cognitive nativist (CN), on the other hand, believes that experience can at most trigger knowledge or concepts that are somehow inborn in the cognizer. (Again, we can talk of more or less extreme cases of cognitive nativism: an extreme nativist will hold that we are born with innate knowledge or concepts, and that experience is irrelevant to what concepts we have; a moderate nativist will admit that experience is needed to activate the innate knowledge or concepts.) Let us now consider the possible reasoning behind the FT interpretation.

(i) The First Reason: The FT is necessary for MR?

It is widely agreed that Aquinas is a metaphysical realist in both of the senses described above and also that it would be a gross mistake not to present him as such. MR is crucial for Aquinas. Aquinas believes that the universe has been created and is continuously sustained in existence by God. God created the world, and in this world, various

---

39 Unless specified otherwise, I use the term 'knowledge' in a non-formal sense of what a given creature believes.

40 What I present here are only simplified and more or less idealized definitions of each view. The features that I ascribe to each position are somewhat exaggerated in order to emphasize differences between them. I am going to expand on the definitions of cognitive empiricism and nativism as I go along explaining Aquinas's and Fodor's positions about cognition.
creatures that are capable, to a greater or lesser extent, of cognition. Metaphysical realism, for Aquinas, is the only position which puts the human cognizers in their right place—as creatures of the divine Creator they can only strive in their cognition to achieve correspondence between their minds and the world. They do not create reality. More on this below

And so, it may seem attractive to interpret Aquinas's views on cognition as in fact *implying* MR. This is, I think, how various Aquinas scholars perceive the FT interpretation. It seems that the FT model could, in principle, be perceived as a secure path towards metaphysical realism that is so crucial for Aquinas. This would be the case were we to interpret the FT as an expression of the special kind of cognitive empiricism. Recall that on the FT account, one and the same form (either numerically or formally the same) appears not only as a sensible species in the external senses (Fs), as a phantasm in the internal senses (Fp), and as a concept in the intellect (Fi), but also as the form of the material thing in the world: $Fs = Fp = Fi = Fo$.

So, the FT model seems to express an even more substantial kind of cognitive empiricism than is usually ascribed to Aquinas. This empiricism is not limited to the relationship between the intellectual and sensory vehicles of cognition, but rather reaches further, to the things in the external world. The empiricist slogan on the FT interpretation should therefore be modified. Instead of saying that "there is nothing in the intellect that was not previously in the senses," it should claim that "there is nothing in the intellect that was not previously in the external world." And so, if the intellect
cognizes things in the world, it seems that these things must exist (or at least must have existed). MR seems to be implied by the FT’s CE.

The attraction of the FT interpretation would therefore lie, at least in part, in the fact that it expresses this special kind of CE which in turn brings with itself the promise of MR. This way the FT, if successful, kills two birds with one stone: it provides an account of the process of cognition and at the same time it implies MR. This seems to be the view, for instance, of Jacobs and Zeis (1997), who explain:

Q16. The interpretation of cognition as actualization of form is supportive of epistemological and metaphysical realism in that there is an intrinsic likeness of cognition and object, on account of what is in the mind being there through transmission of form. (Jacobs and Zeis, 1997, p. 541)\(^{41}\)

Jacobs and Zeis identify the whole of the Aristotelian-Thomistic theory of knowledge (they call it the A/T theory) as a hybrid theory that is at the same time metaphysical and epistemological. They say:

Q17. The A/T theory of knowledge is not an epistemological theory (…), if by ‘epistemological’ we mean one which is clearly demarcated from a metaphysical, psychological, or even semantical theory. The A/T is a metaphysics of knowledge. (Jacobs and Zeis, 1997, p. 539)

The idea that the FT interpretation is also such a hybrid theory, including both epistemological and metaphysical claims, is not a problem in itself. It would become problematic, however, were one to think that its metaphysical claims require its

\(^{41}\) Cf. also Panaccio: "What more direct form of realism could one hope for than such a doctrine which says that the very nature of the external thing—its essence—comes to exist in some way within the cognizing subject?" (Panaccio, 2001, p. 187).
epistemological claims, i.e., if one was tempted to think that MR is indeed dependent on CE or vice versa. As we shall see, the FT in its first two versions is not necessary in order to preserve Aquinas's MR; we could deny Aquinas's cognitive empiricism without any threat against his MR. This is because (1) cognitive empiricism does not imply metaphysical realism, and (2) cognitive nativism does not imply metaphysical idealism. Rather, there is really no relationship of dependence between CE and MR. In addition, (3) it turns out that it is not a warranted claim that the FT is actually an expression of a strong version of CE. On the contrary, it seems that it easily allows for both a CN and a MI readings of Aquinas.

(ii) The First Reason is No Good

Consider the claim to the effect that CE is necessary for MR:

\[ C^*. \neg CE \rightarrow \neg MR \]

(given that CE = \neg CN and that MR = \neg CI, C* is equivalent to C**. CN → MI)

C* says that CE is a necessary condition for MR. It suggests that if Aquinas is not an empiricist about cognition, he couldn't be a metaphysical realist either, or that if Aquinas is a cognitive nativist, he must also be a metaphysical idealist.

If cognitive nativism was indeed the case, that is, if our concepts, i.e., forms in the mind (Fi), are prior to experience, then it seems that we may not need experience to acquire them. (Of course this will also depend on the definition of innateness that we accept.) And if no experience whatsoever is needed, then neither is any kind of
interaction with the external world. The external world could as well not exist, or be
nothing but a creation of our own minds, and we would still have the same concepts. In
such a case the existence of concepts would be no guarantee that concepts represent
anything in the world. If concepts are innate, that is, already there in the mind before we
actually cognize anything, and if, in addition, Fi = Fo as the FT suggests, then it also
becomes a possibility either that we cognize only our own mental states, or that the
world that we see is structured by our cognitive apparatus. And so, it seems that CN
easily allows for MI (at least when CN is added to the FT).

At the same time, however, CN on its own certainly does not imply MI. Being a
nativist with respect to concept and knowledge acquisition does not preclude remaining
a metaphysical realist. Think, for instance, of two most radical cognitive nativists, Plato
and Fodor. For Plato, knowledge does not come from experience, but is only triggered
by it,42 and for Fodor, concepts are not produced in the process of cognition but "are
already there, waiting to be triggered" (Fodor, 1998, p. 129). However, their position on
the acquisition of concepts or knowledge is still consistent with metaphysical realism,
with the belief that there exists a reality that is independent of any cognizers. (For Plato
this independent reality is the platonic Heaven inhabited by Forms; it is completely
independent of any cognizers, but we do have access to it—well, at least some of us
philosophers do.) So, CN is not a sufficient condition for MI (¬(CN → MI)).

42 Cf. Plato, *Meno* 82–85; *Phaedo* 73–77a. In Aquinas's words, for Plato "naturally man's intellect is filled with
all intelligible species" (*ST* I 84, 3; cf. also *ST* I 117, 1 and *DV* 10, 6).
As for CE, it says that our concepts are not innate but come from experience. The concept CAT, for instance, would most likely be acquired from experience with cats. And so, if I indeed possess the concept CAT, then it seems that at least one cat must have existed, and that MR must be the case. Well, things are unfortunately not so simple. Metaphysical idealists also believe that we have experiences with cats. They could agree that we get our concept CAT from our experiences with cats. On their view, however, cats do not have an independent existence; they are, rather, constructs of our minds. (And even for the realist, my experience with vampires—in my dream—does not prove the independent existence of creatures that drink human blood.) And so, as it was the case with the supposed metaphysical implications of CN, so it is with CE. There aren't any such implications. On its own CE is independent of both metaphysical realism and idealism. (This is confirmed by the fact that among philosophers we can find both CE-sts, like Locke, who are MR-sts, and CE-sts, such as Berkeley, who are metaphysical idealists.) So, CE is not a necessary condition for MR (~(~CE → ~MR)). This refutes claim C*. We may conclude, therefore, that one could safely reject the FT with the kind of CE that it supposedly expresses without any threat to Aquinas's MR.

(iii) Second Reason for the FT

The second possible reason for Aquinas scholars to favor the FT is close to the first: it stems from their desire to preserve Aquinas from the threat of skepticism. While Aquinas indeed believes that there is an objective world independent of our perceptions, and that human beings can know about the nature of this world, he also talks of
intelligible and sensible species as the means of cognition, and as similitudes or representations of objects cognized. This raises the question whether in his views on cognition he is a direct or rather an indirect realist.\textsuperscript{43} If he is an indirect realist, skepticism becomes a serious threat.

For a direct realist, cognition occurs in virtue of a direct access to things in the world. On this view, we are directly aware of real, physical objects. Perception, which is taken to be a reliable source of knowledge, both proves that and explains how we know about the nature of the objects in the external world. On the other hand, indirect realists, or representationalists, emphasize that in perception we are directly aware only of our own mental states (images, ideas or representations) that are supposed to represent things in the external world. Representationalists believe, therefore, that our access to the external world is mediated. Metaphysical realism is not obvious on such a view. One could assume that there exist real objects corresponding to our mental images, because this, supposedly, would be the best explanation why we actually have the sort of mental images we do. But skepticism is an option here as well. It may turn out to be a more plausible position that we actually cannot know that there is an objective world, even if this is the case, or that we cannot know what this world is like; we cannot know whether our mental states correspond to anything outside of us, to any external world.

So, for instance, Locke holds that all we are directly aware of are our ideas. If this is the case, then we cannot really know whether (or how) they correspond to the world

\textsuperscript{43} That Aquinas is an indirect realist, or a representationalist is claimed, for instance, by Paul Hoffman (1983), Sharon Kaye (2006), and Dominic Perler (2000). Cf. also Panaccio, 2001, for a discussion on how the two positions, direct realism and representationalism, are in a way both present in Aquinas.
outside. And even if we do manage to somehow resolve these issues, there still remains the question whether we can know anything about the nature of that world. It may be that the only thing that we do know is that this world is the cause of our ideas.\textsuperscript{44}

The worry about skepticism supposedly threatening Aquinas's account of cognition could be explained as follows: if Aquinas is a representationalist about cognition, then, on his account, we have something like a veil of ideas between the cognizer and the world cognized. Because of it we no longer know whether what we have in the mind represents what's really in the world—we cannot know whether there is any external world in the first place. All we can know is that we have certain representations or images in the mind. And so, if such was Aquinas's account of cognition, then it would not be obvious whether it was still legitimate for him to hold on to his metaphysical realism.\textsuperscript{45} From this perspective, the role of the FT interpretation would be either to present Aquinas as a direct realist, or to show that the version of representationalism that is implied by the FT account of cognition is not threatened by skepticism.

The Numerical Identity version of the FT seems to be the most resistant to the skeptical challenges, and so it is probably the best bet for those who want to claim that Aquinas is a direct realist about cognition.\textsuperscript{46} On the Numeral Identity theory, what's in

\textsuperscript{44} Locke believes that we have at least some probable knowledge about the nature of that world. It is not clear, however, that he is entitled to believe that.

\textsuperscript{45} This is a familiar objection to Locke: it seems that he cannot both present his representational account of cognition and disavow skepticism about a mind-independent physical world.

\textsuperscript{46} The Numerical Identity version of the FT is not the only option to avoid skepticism. Another way in which this can be done may seem surprising: it is to actually assume metaphysical idealism, according to which there is just no such thing as an external objective world. All that there is, is the mind and the images.
the mind is the actual form of the object, and not its representation. If so, then the presence of this form in the mind seems to be the proof of an interaction with an independently existing external object. For the form of the object to be in the mind, it seems that the object itself had to exist.\footnote{Notice, however, that on the FT, when we think about horses, that is, when we intentionally exemplify the form or property of being a horse, all that this implies is that this very form or property exists. In order to avoid skepticism, the FT would also have to claim that intentional exemplification can occur only if there was prior natural exemplification (cf. Haldane, 1998, p. 269). It is not clear, however, that this is Aquinas’s own claim or that he could provide any plausible support for it.} On the Identity version of FT, there is just \textit{no such thing} as a veil of ideas between the cognizer and the object she cognizes.\footnote{This being said, it seems that one can still raise skeptical worries even without a veil of ideas doctrine once one allows for states of illusion phenomenologically indistinguishable from those in which we actually receive the identical form.}

Finally, the proponents of the FT interpretation may try to find support for their view also in the fact that, as we said (above in 1.1.2 (v)), Saint Thomas talks about intellectual cognition as consisting in what is called the process of \textit{abstraction}. Aquinas explains:

\begin{quote}
Q18. to abstract the universal from the particular, or an intelligible species from phantasms [means] to consider the nature of the species without considering the individual principles that are represented by the phantasms. (\textit{ST} I 85, 1, ad1)
\end{quote}

In the process of intellectual cognition, Aquinas similarly explains in \textit{De Veritate}, "the likeness of a thing existing in our intellect," that is, the intelligible species, is abstracted from a phantasm and received in the intellect "separated from matter and all the conditions of matter, which are the principles of individuation" (\textit{DV} 2, 6). Abstraction
consists in the intellect's looking at the phantasm and focusing on certain relevant (essential) features of the object of cognition while ignoring all others. If the intellect is supposed to focus this way on some features of the object, those features must be there already present in the phantasm (more on abstraction below, 1.2.3 (vi)).

The FT model interprets both passages above (from *ST* I 85, 1, ad1 and *DV* 2, 6) as suggesting that phantasms and intelligible species are indeed *the same forms*, or, on the Formal Sameness version, that they are different kinds of mental items which represent the same forms; they are the same likenesses of the object cognized. It seems that this indeed is the only way in which abstraction could work: on the assumption of sameness of forms. What's abstracted on a higher level has to be there, already present, on a lower level.

As it turns out, however, the FT on its own is unable to play the role that the above considerations try to ascribe to it. In the last part of 1.2.3 we are going to show that the interpretation of cognition in terms of *abstraction* as it is commonly understood by the supporters of the FT interpretation needs to be rejected.
1.2.3 Problems with the Formal Sameness version of the FT interpretation

(i) The FT model allows for both CN and MI

The main claim of the FT is that the form informing the object (Fo), the form in the senses (Fs), and the form in the intellect (Fi) are the same: Fo = Fs = Fi. Identity, however, is a symmetrical relation. (Formal sameness, as type-identity, is also symmetrical.) Forms in the mind may indeed represent forms of objects cognized, but because the relationship between them is symmetrical, forms of objects could as well represent forms in the mind. If all that the FT says is that the forms involved in the process of cognition are the same (either numerically or formally), then this on its own guarantees neither MR nor CE. If we agree, as the FT suggests, that concepts (Fi) are exactly the same (token- or type-identical) as the original forms informing objects in the external world (Fo), then what we get is the claim that the structure of things in the world matches the structure of concepts in the cognizer's mind (Fi = Fo). But this does not preclude the possibility that forms Fi are prior to forms Fo and Fs. Forms in the intellect may exist in the cognizer's minds prior to any experiences that she ever had. If experience is not needed for concept acquisition, then CE is negated.

49 On the other hand, the relation of similarity of forms that I propose with the FTF account of cognition is not symmetrical. Cf. 2.2 (ii).
The FT seems to be a position which easily allows for metaphysical idealism. If the concept X is just a copy of the object X, then the content (and the structure) of the concept already indicates the structure of the thing—more generally, the content of concepts indicates the structure of the world. But this is consistent with the view that there is no world that would be independent of our cognition. What we call the 'external world' could be nothing but a creation of our own minds. Thus, the FT on its own seems to allow for MI as at least a serious possibility.

(ii) FT Does not Guarantee MR; something else does

The FT interpretation on its own does not defend Aquinas from the threat of MI. It is not enough to claim, as the FT interpretation does, that all forms involved in the process of cognition are the same, and that there is the same structure between the mind and the world. What's crucial is to show that it is the structure represented in the mind that comes from the world, and not the other way round. The FT emphasizes the sameness of forms. Sameness of forms, however, is not enough. On the contrary, sameness of forms is itself rather problematic (see 1.2.3 (iv) and (v)).

In addition, it turns out that there is no need for sameness of forms in any of its version to prevent Aquinas from both MI and skepticism. Rather, it seems that certain elements already present in Aquinas's account of cognition prevent it from being threatened by MI and skepticism. For one thing, even though Aquinas does hold that intelligible species are representations, he also believes that they are not what we cognize, but that by which an object is cognized (ST I 85, 2). The same applies to Aquinas's
understanding of sensible species (and phantasms). On Aquinas's explanation of
sensation, when you eat a candy and taste its sweetness, your sense of taste receives the
sensible species 'sweet.' But Aquinas doesn't mean by this that there appears a sense
datum somewhere in your body separate from the sweetness of the candy. Rather, your
taste's tasting of sweet and the sweetness of the candy are one and the same thing. As
Aquinas puts it, sense in act is the sensible in act (ST I 87, 1, ad3). And so, on Aquinas's
account, there is in fact no such thing as a veil of ideas between the cognizer and the
object she cognizes. Since the senses are passive, if the form of an object is 'stamped'
upon them, then the object must have existed. The actualization of a sense organ is the
same thing as the activity of a sensible form of an object. Aquinas is of course aware that
sometimes we may dream or hallucinate various sensory experiences. We may then only
think that we see something; in that case, it is 'as if' the form of an object was imposed on
our sense of vision, even though there was no external object there at all. Aquinas
wouldn't consider this to be a problem for his metaphysical realism. He believes that the
external senses, even though they may sometimes malfunction, are for the most part
reliable; they were created by God in order to allow us to successfully function in the
external world.

A similar explanation as we have seen with respect to the senses and their objects
applies also to intelligible species. As John O'Callaghan's explains, a concept, for
Aquinas, is nothing but a modification of the intellect, nothing but "the informed activity
of the intellect as it grasps res extra animam" (O'Callaghan, 2003, p. 168). Thus, just as in
grasping a pen in order to write "there is no third thing that exists between my hand and
pen (...), [s]imilarly, there is no third thing other than the conceiving intellect and the res extra animam" (O'Callaghan, 2003, pp. 169–170). The intellect (primarily) cognizes things, and not its own mental states.\textsuperscript{50}

Aquinas himself seems to explicitly deny representationalism. In De Veritate he says:

Q19. What is cognized by intellectual sight are the things themselves, not their images (imagines). (...) the object of the intellect is the very essence of the thing, although the intellect cognizes the essence of the thing through its likeness, as through a means of cognizing, and not as through an object over which its vision is first carried. (DV 10, 4, ad1; cf. also DV 3, 6 and DV 10, 6, obj. 1)

What the intellect cognizes are things in the external world, and not representations of these things produced by the intellect. Consider also Gilson's explanation:

Q20. The sensible is the union of a form, and hence an intelligible, with determined matter. Therefore, the sensible contains in potency some intelligible (...). If we pass on to man we find something intelligible in act, his intellect (...). But we know that this intelligible lacks determination. It is a light by which we can still see, but (...) if it is to enable us to see, this light must fall upon some objects. But for it to fall upon objects, there must exist objects related to it. The intelligible in act, our intellect, will die of inanition unless it finds nourishment in the world in which we are placed. (Gilson, 1961, pp. 219–220)

Aquinas himself compares the intellect, that is, the active power of understanding, to light. We can only see a thing if light falls upon it. If there is no object that would reflect light that falls upon it, then there is no seeing. Similarly, we can understand a thing only if the intellect 'falls upon it'; that is, if the intellect finds something intelligible in it. For

\textsuperscript{50} Aquinas should be classified as a representationalist if we define representationalism as "a theory of cognition which attributes a crucial and indispensable role to some sort of mental representation" (Panaccio, 2001, p. 185). Mental representation is then understood as "any symbolic tokening existing in some individual mind and endowed within this mind with a semantic content" (Panaccio, 2001, p. 185). On Aquinas's account mental representations are the means and not objects of cognition.
this reason, the object that the intellect cognizes also has to exist, i.e., to have being (see below, 3.2).

The intellect can also create some of its objects of cognition, but this kind of intellectual activity is only secondary on Aquinas's account. Also, ideas or concepts may become a secondary object of cognition, as when the intellect reflects on itself. Aquinas calls the way in which the intellect grasps its own actions 'a kind of return' or 'a kind of reflection' (DV 10, 9) and he compares it to the way in which we look in a mirror:

Q21. Through a likeness within vision obtained from a mirror, one's vision is directly drawn to cognize the reflected object; but by turning back in a way it is drawn through that same [internal likeness] to the likeness that is in the mirror. (DV 2, 6)

Sensory cognition of individual physical things existing in the external world, cognition resulting in the production of phantasms, is a necessary condition, and a basis, for intellectual cognition ever to take place.

Aquinas's greatest concern is with truth which he explains as an agreement of intellect with things. As he says in his Summa Theologiae, "truth is defined as an agreement between the intellect and the thing" [per conformitatem intellectus et rei, veritas definitur] (ST I 16, 2, ad2; cf. also DV 1, 1-2 and ST I 84, 2; on cognition of truth see DV 1, 9). To acquire truth means for the intellect to adjust itself to how the world really is. The ideas of the intellect are of any use only if they are measured by things. And again, for all this to be possible, there must exist a world of things independent of our intellect and accessible to our senses. Thus, metaphysical realism is a must.
At this point, one might object that metaphysical realism does not follow just yet. The theory of truth discussed so far claims that what is real is independent of my particular mental states, and not that reality is mind-independent. In principle, then, it could still be the case that the world is somehow constituted by our senses. If so, Aquinas could keep his idea that truth consists in adjusting the intellect to how the world really is, despite being a metaphysical idealist in holding that the world that the intellect adjusts itself to would be a creation of the senses (in such a case Aquinas would hold a view similar to Berkeley's).

This objection, however, does not work against Aquinas. Aquinas clearly distinguishes two following kinds of cognition. In God, cognition has a top-down direction. It consists in the agreement of things to God's intellect. God's thinking is really equivalent to His creating things in the world. On the other hand, in all things created by God, cognition goes the other way round—it has a bottom-up direction, and it consists in the agreement of the intellect (or in general, of cognitive faculties) to things in the world. (This kind of cognition, cognition by assimilation to (or by means of copying the form of) things, and not cognition as creation of things, is what shows the place of human (or other kinds of) cognizers as different from the divine cognizer-Creator.)

For Aquinas, intellectual cognition for humans is possible when proper and incidental sensibles somehow give rise to intelligible species. It is not the case, as metaphysical idealism would suggest, that intelligible species are pushed down and 'stamped onto the world.' On the contrary, for Aquinas "knowledge in us is the stamping of things on our minds" (Haldane, 1983, p. 235). Cognition goes only in one
direction: it consists in taking the forms from objects in the world, and transforming
them into forms in the cognitive faculties.

It is for Aquinas the main feature of the external senses that they are entirely
passive. (There is no need for the senses to be active because things in the world, or their
individual properties, are actually sensible.) The external senses are like thermometers:
they only register stimuli incoming from the external world. And, as we would not want
to claim that temperature outside is somehow constituted by the thermometer that
measures it, similarly it does not make sense to suggest that the sound of thunder is
constituted by the sense of hearing. We can only talk of the sense of hearing if there
exists some kind of stimulus that activates it. If there is no such stimulus, then there is no
hearing either; all that we have is an oddly constructed piece of meat attached to one's
head. The functioning of the external senses, on Aquinas's account, proves the existence
of the external world.51

As I mentioned earlier, even though the concerns regarding MI or skepticism are
likely motivations of Aquinas's interpreters in their support for the FT interpretation, it
is highly unlikely that Aquinas himself would have similar concerns. If he talks about the
identity of forms it is probably because this is what he finds in Aristotle, for instance in
his De Anima. The fact that Aquinas uses the same terms as Aristotle, however, does not

51 Cf. again Aquinas's idea that sense in act is the sensible in act, for instance in ST I 87, 1, ad3. This idea is
present in Aristotle: "The act of the sense-object and the sensation are one and the same" (DA III 2, 425b26;
different translation in McKeon (1947): "The activity of the sensible object and that of the percipient sense is
one and the same activity, and yet the distinction between their being remains."). The activity of an external
sense proves the presence of the sensible acting upon it—they are one and the same thing. But they are also
distinct in that the 'sensible in act' is an accident of an externally existing physical thing, and 'sense in act' is
a modification of cognitive faculties of the cognizing subject.
impose a literal understanding of them. Literal understanding, as we shall see, would make Aquinas's position quite implausible. But Aquinas has no serious reason to keep it. In addition, various passages in Aquinas suggest that he had in mind a different and more plausible interpretation.

Aristotle himself wasn't worried about skepticism or about real existence of the external world. Skepticism became a threat for early Christian philosophers (cf. Pasnau, 2002, pp. 213–4). Augustine, for instance, was seriously concerned with skeptical problems (cf. his *Confessions*, esp. VI 1, 2, 4, 5). During Aquinas's times, however, skepticism was simply not in the air. As Robert Pasnau explains:

Q22. Later thinkers seem to have regarded Augustine's treatment of [skeptical] issues as decisive. Skepticism simply ceased to be a prominent topic of discussion until the end of the Middle Ages. Instead, attention was focused on how knowledge was acquired. (Pasnau, 2002, p. 214)

Skepticism became a widely debated issue only in the post-Cartesian world. This is most likely why contemporary Aquinas scholars are concerned with it: they are, after all, post-Cartesian. The first two versions of the FT interpretation are considered the most common reading of Aquinas. There are, as I am going to show, important reasons why they should be rejected. Still, it is worthwhile, I believe, to be aware of possible motivations behind the FT, even though they are not Aquinas's own. In addition, to the contemporary reader the idea of the involvement of form in the process of cognition, and the talk about formal identity and transmission of identical forms must seem really foreign. The discussion of reasons behind the FT interpretation broadly understood
shows what can be still attractive in what at first seems an outdated and obscure theory to contemporary philosophers dealing with currently discussed issues.

Our conclusion at this point is, then, that

C1. The FT *neither guarantees nor is necessary* in order to preserve Aquinas's metaphysical realism.

One may be a metaphysical realist without commitment to the FT interpretation or indeed to any particular theory of cognition. What we are going to do next is to show that:

C2. The FT has certain consequences that Aquinas would certainly want to avoid (parts (iii) and (iv) in 1.2.3).

C3. Given certain important passages from Aquinas's writings, the interpretation of his position proposed by the FT model turns out to be wrong (part (v) in 1.2.3).

(iii) Cats in the mind

The problem with the FT that we are now going to consider concerns not so much the issue of *sameness* of form, but rather the very notion of *form*. Recall that for Aquinas, 'form' is explained as that which makes something be what it is. The form 'cat,' therefore, is what makes something a cat. Concepts are supposed to be forms in the mind. But if the form 'cat' is what makes something a cat, and if in my head I have such a form, why, then, doesn't my concept CAT produce a cat in my head? It seems that if what I have in my mind is the concept-form CAT, then I must have an actual cat in there as well. Forms, for Aquinas, cannot exist on their own: "forms cannot exist before their matter" (*InDA* III 10, 743). And so, if there is a cat-form in my head, there also has to be some cat-
matter in there as well. It seems, therefore, that I must have a real cat in my head if I have the CAT concept, and that possessing the concept CAT in my head is really equivalent to having in it a cat itself. If we want to keep the idea that concepts (as well as other vehicles of cognition) are forms, then in order to avoid problems of the 'Cats in the mind' kind, we have to figure out what kind of forms are those forms that we find in the mind. Also, we need to try to determine whether, and if yes, how, forms in the mind actually inform the mind and still play a form function.

Notice, first of all, that the 'Cats-in-the-mind' problem arises only if we mistakenly identify either the form with the thing, or the form with the concept. What causes this mistake is the definition of form as 'what makes something be what it is.' A more correct definition of form, however, is rather that together with appropriate matter it determines something to be the kind of thing it is. Forms, for Aquinas, are metaphysical principles. Anything that exists, exists in virtue of the form that it has: "everything exists through the form" (DP 3, 1). Also, anything that exists is the kind of thing it is because of the form that it has. But forms can perform their function only when they in-form some kind of matter. This suggests indeed that forms as such cannot exist 'on their own.' It is possible to think of forms 'on their own,' not informing any kind of matter, but that's the only way in which completely de-materialized forms can exist in the world: in our thoughts.

Consider the following passages:

---
52 This problem concerns all metaphysical and epistemological accounts relying on the matter/form distinction.
53 When we consider any individual physical object, we can focus on one of the metaphysical principles constituting it, i.e., its form. But every individual thing is a compound of form and matter. On the other hand, de-materialized forms can and do exist in the supernatural world. Aquinas explains that only "in
Q23. This also should be considered, that things existing concretely in Nature—physical things—are not alone in having their essences in matter; the same is also true of mathematical entities. For there are two kinds of matter: sensible matter, which is intrinsic to physical things and from which the mathematician abstracts; and intelligible matter, intrinsic to mathematical entities. (*InDA* III 8, 707)

Q24. For there are some forms which can only exist in a matter which is possessed of certain definite sensible qualities; and such are the forms of physical things; and such things therefore always involve sensible matter. But there are other forms which do not call for matter possessed of definite sensible qualities, yet do require matter existing as quantity. These are the so-called mathematical objects such as triangles, squares and the like; they are abstracted from sensible matter, but not from intelligible matter (...). Clearly then, both physical and mathematical objects have their forms in matter, and in both there is a difference between a thing and its essence; which is why in both cases many individual things are found to share the same nature: e.g. men and triangles. (*InDA* III 8, 708)

All forms that human beings encounter in the world that they inhabit inform some kind of matter. Even abstract mathematical objects are constituted by forms in-forming what Aquinas calls the intelligible kind of matter. This applies also to forms in the mind. Forms need to be implemented in matter to make a thing be what it is, a form/matter compound; you cannot have the form without matter, or vice versa. The form 'cat' on its own does not make anything a cat. In order to produce a living cat a certain amount of secondary matter is needed. But obviously, cat-making secondary matter is not in the cognizer's head, even when its form is. So when a cat-form appears in the mind as a concept, a real cat cannot be obtained this way. All we can get is a cat-representation. In perfectly immaterial substances the thing is identical with its essence" (the thing equals its form), but "such substances are beyond the reach of the human mind" (*InDA* III 8, 710).

54 Consider another example. On Aquinas's account, if we want to explain what it means to be a house, we'd have to say that it is to have a house-nature, not just a house-form. Even though these terms are often used interchangeably, 'nature' is more than 'form.' Form is not enough to make a house; rather, it belongs to the nature of the house that some matter is needed in order for it to exist. In fact, matter of a specific kind is needed (e.g., a house can be made of brick, or of wood, but not of water) in which the form will be implemented. (Within the set of all things constituted partly by the form 'house,' we could further distinguish various subsets. We'd say that some houses are constituted more concretely by the form 'brick-house,' some by the form 'wooden house,' etc.)
the mind, the form 'cat' is what makes a mental representation be the representation of a cat. The form as a concept is implemented into some kind of mind-matter.

The term 'mind-matter' may at first seem completely counterintuitive. After all, for Aquinas the crucial feature of the mind is its immateriality! The reason why it sounds so odd, however, is that we are not used to the medieval vocabulary. In that vocabulary, matter does not necessarily imply something physical. Rather, at the most fundamental level, 'matter' refers to something potential. And so, the only requirement for us to be able to talk about mind-matter is that there is in the mind some potentiality. And indeed, on Aquinas's account, the intellect (as such, that is, independently of its current activity) can plausibly be understood as pure potentiality, as a capacity to acquire, or to be in-formed by forms of things.

Both intelligible matter and mind matter, therefore, can also be described as kinds of potentiality.

Q25. In any nature which alternates between potency and actuality we must posit (1) a factor akin to the matter which, in any given class of things, is potentially all the particulars included in the class; and (2) another factor which operates as an active and productive cause, like art with respect to its material. Since then the intellectual part of the soul alternates between potency and act, it must include these two distinct principles: first, a potentiality within which all intelligible concepts can be actualized (this is the potential intellect) (...), and then, also, a principle whose function it is to actualize those concepts. And this latter is the agent intellect (InDA III 10, 728)

Aquinas explains that the potential intellect is "free from matter and impassible and pure" (InDA III 7, 677-83). In his terminology, the potential intellect in itself is pure potentiality, and pure potentiality means exactly the potentiality to acquire any form at all. It is for this reason that the intellect can be actualized by all kinds of different concepts;
every different concept actualizes this potentiality in a different way. Since only matter can receive form, pure potentiality, therefore, turns out to be the same thing as pure materiality.

To be more precise we should say that it is only God’s intellect that can be understood as pure potentiality. God’s intellect is truly unlimited. God can cognize anything he could possibly want to cognize. The human cognitive capacity, on the other hand, is clearly not unlimited in this way. Most likely there are things that our intellect just cannot think. Similarly as the sense of taste can be described as a ‘taste-matter,’ that is, as potentiality to receive taste information, also the human mind can be described as ‘(human) mind-matter,’ that is, a potentiality limited to in-formation by what’s intelligible for human beings. (In contemporary terms, this potentiality of the intellect can be understood as the potentiality of our brain for thinking different thoughts and thought-parts. Here again it makes sense to say that there are some things that the human brain just could not think. See below, 2.3.4 (iv). See also 5.4.4 (i).)

Another reason why we need to posit mind-matter is in order to be able to explain how our minds can have different concepts. We have different concepts because of their different objects. But in addition, there has to be something in my head that distinguishes my concept CAT from my concept DOG. On Aquinas’s account, we’d say that we have different concepts when different 'chunks' of mind-matter are informed in a different way.

Cf. the following passage: "The intellect (...) has one natural object, of which it has knowledge per se and naturally. And this object must be that under which are comprised all things known by the intellect: just as under color are comprised all colors, which are per se visible" (SCG II 83, 1678).
To sum up, for Aquinas, a form on its own does not make anything be what it is. A form on its own is just a form; it is a metaphysical principle that does not exist without matter that it informs. It determines some thing to be the kind of thing it is: "Forms bring a certain identity to things, causing them to be units of a certain kind. This is no less true of concepts than it is of res extra animam" (O'Callaghan, 2003, p. 238). The form of a cat is not yet the cat, and the form 'cat' in the mind is not yet the concept CAT. To have a cat, in the world or in the mind, we need both form and matter. A certain chunk of secondary matter must be informed by the cat-form to make it into a cat. When the form 'cat' informs mind-matter, we have a CAT concept. The form 'cat' in my head is not a cat, or rather, my mind is not a cat, for the same reason for which the blueprint of a building is not itself a building. The blueprint is a physical representation of the form of the building put on a piece of paper; it is the configuration of the elements of the building encoded in a schematic way. In Aquinas's terms, the same form that exists materially as the form of the building has an intentional existence in the blueprint. The building's form is its constitutive principle; it is the encoding of the information about its structure, etc., as it is already realized in the brick matter. When we talk about the blueprint, the same form has an intentional existence: the information about the structure of the building is encoded on paper. The drawing on the piece of paper represents the structure of the building in the sense that it contains the same information (cf. above, 1.1.3 (i) on forms as containing information).

To be precise, Aquinas would never claim that a blueprint has the form of the building, but rather only that it represents the form of the building. Only the former claim
would raise the question: "Why then isn't the blueprint the building?" The second claim raises the question: "What is it for the blueprint to represent the form of an actual building?" It does not mean to have the same form as the building. Rather, it means to contain the same information, but to encode it in a different way and in a different medium.

The role of form depends, in the end, on the kind of matter that it informs. So, the answer to the Cats-in-the-mind problem is that depending on what kind of matter a form informs, it may either constitute a physical being in the external world, or an intentional being in the intellect. Cats-in-the-mind, therefore, turns out not to be a problem for Aquinas. What constitutes a much more challenging issue is the FT's claim about sameness of forms at every level of cognition. If this was indeed the case, that is, if all the forms involved in the process of cognition were exactly the same, then it seems that Aquinas should not really talk of three distinct levels of cognition. In addition, a careful analysis of various passages, especially from *De Veritate*, will show as pretty much obvious that for Aquinas the forms at different levels of cognition cannot be the same. (The second problem will also be discussed in 2.1.)

(iv) No distinct formal objects for different levels of cognition?

The most serious problem with the FT concerns its claim to the effect that all forms involved in the process of cognition are either numerically or formally the same. As we said, on the FT, intelligible species are nothing but phantasms which have been stripped of their individualizing conditions: they are distinct mental items but they represent the
same form or the same aspect of the object cognized. When we cognize a cat, the intelligible species represents the nature CATNESS. On the first two versions of the FT, the same form, the nature CATNESS, would also have to be represented by the phantasm.

For Aquinas, it is a legitimate move to distinguish different cognitive faculties, only if each of them has its own formal object. But then, because of the fact that on the FT cognitive faculties at each stage of the cognitive process operate on the same forms, it may seem that Aquinas's distinction of different cognitive levels is not justifiable. If the forms involved in the process of cognition are not formally distinct, if they can't be considered distinct formal objects, then there is no ground for Aquinas to distinguish between different levels of cognition in the first place.

The question that we need to ask now, therefore, is whether the FT's claim that all vehicles of cognition are really the same forms would indeed make it impossible to still hold that there are different formal objects for different levels of cognition. Well, for one thing, one might wonder whether it is a legitimate move to equate the 'formal object of the cognitive power' with 'form.' One could try to defend the FT and argue, for instance, that even though it is indeed one and the same form that appears in all the stages of the cognitive process, still at each stage this same form occurs with a different modality. And that different modalities with which the forms occur are sufficient to allow us to distinguish different formal objects of different cognitive faculties. So, for

---

56 Cf. ST I 77, 3: "Wherefore we seek to know the nature of a power from the act to which it is directed, and consequently the nature of a power is diversified, as the nature of the act is diversified. Now the nature of an act is diversified according to the various natures of the objects."
instance, even though it is the same form that at some point appears as a phantasm, and
at a later point as a corresponding intelligible species, Aquinas could still say that the
formal object of the intellect is intelligible species, that is, the form of the object stripped
of individuating conditions, and that the formal object of the senses is the same form
when it is to some degree en-mattered (cf. ST I 85, 1: "The object of every sensitive power
is a form as existing in corporeal matter."). On both levels we would have the same
content, but existing in two different media. Phantasms and intelligible species would
differ as different kinds of vehicles of cognition, and so, they could be seen as different
formal objects, but they would count as the same form because of what they represent.

The difference of modality between phantasms and concepts is obvious: the
former informs sensory matter, the latter—mind-matter. As existing objects (as mental
items in the mind), phantasms and intelligible species are different, and this is just
unquestionable (for any two objects, if they are of different kinds, this is because they
have different forms). What is questionable, however, is whether the FT, which claims
that the vehicles of cognition at different stages of the cognitive process are not distinct
in what concerns their role of representing, can still be held as an acceptable position
allowing us to account for different formal objects for distinct levels of cognition.

Let us now try to clarify what Aquinas means by a formal object. Aquinas
distinguishes between things in their own autonomous reality, and the same things
under the relative aspect of their attainability by some cognitive power. Any given thing
can be described from two perspectives: (a) as a material object, i.e., when the thing is
described as it is in itself; and (b) as a formal object, i.e., when it is considered as an
object of an activity, an object of the exercise of a power. In the second case we deal with
the object as it is specified by some act. For instance, we can consider my friend John in
himself, as a material object: from this point of view he is a human being, a man. On the
other hand, depending on what perspective we take, different descriptions can be
applied to John considered as a formal object: he can be described as something visible,
as somebody's son or father, as a law enforcement officer, etc. 57

For Aquinas, the formal object of a given faculty is that object which makes the
faculty (or its operations) of the kind it is. Every faculty or power has a formal object,
which defines it as the kind of power it is. 'Formal object' is just another name for 'the
object that defines a power as the kind of power it is.' And so, the formal object of the
will is good—this is what defines the nature of the will. The formal object of the power
to walk is walking. We can talk of different faculties, or different powers, only if it is
possible to identify distinct formal object for each of them. And so, when Aquinas says
that, with respect to cognition "diversity in objects known implies diversity in the
knowing faculties" (InDA III 8, 711), what he has in mind is diversity of formal objects of
cognitive faculties.

In order to find out what formal objects our cognitive faculties can possibly have,
we need to consider the main function that these faculties are supposed to have. For

57 Cf. the following passages: "The object of every cognitive habit includes two things: first, that which is
known materially, and is the material object, so to speak, and, secondly, that whereby it is known, which is
the formal aspect of the object. Thus in the science of geometry, the conclusions are what is known
materially, while the formal aspect of the science is the mean of demonstration, through which the
conclusions are known" (ST II–II 1, 1). "The unity of a faculty or habit is to be gauged by its object, not
indeed, in its material aspect, but as regards the precise formality under which it is an object. For example,
man, ass, stone agree in the one precise formality of being colored; and color is the formal object of sight"
(ST I 1, 3). "Forasmuch as nature is ever directed to one thing, it follows that of one power there is naturally
one object, for instance color is the object of sight, sound of hearing" (SCG II 83, 1678).
Aquinas, this is the function of representing. The main function of forms processed by cognitive faculties is to represent. In the context of cognition, the 'form,' which is also called 'species,' or a 'likeness,' is really a synonym for 'representation': "the forms received in cognition are such that when they are received by their subject they make other things (namely, the objects they represent) be present to the subject as well (though only intentionally)" (Brower, 2008, p. 6). The function of representing is what shows that there is indeed a link between cognitive faculties and objects of cognition; it also shows why the cognitive faculties are useful for the cognizing subject in the first place: they inform the subject about the object in the external world.

But then, if representing is the main function of cognitive faculties, then their distinction must be based on what kind of thing each of them represents. If we want to further distinguish different levels of cognition, then there has to be some kind of representation that each level is characterized by.

Individual external senses differ with respect to their proper objects — each receives a different kind of information from the external world, and each produces a sensible species that represents a different accidental feature of the object cognized. At the same time, when we consider the external senses as just external senses we can say that they share the same formal object: they have one function — to produce sensible species that represent distinct physical features of individual things. It is similar with the internal senses: they also have a common function: to represent the same unified percept, the unified whole made up of representations of different accidental features of
the object cognized, produced by the external senses, that are put together by the common sense; but each does it from a somewhat different perspective.

It is true that on Aquinas's account, on every level of cognition the form exists with a different modality, that is, it is de-materialized to a different degree. However, the form's modality is just its accidental feature. It is not sufficient to distinguish different kinds of cognitive faculties, and even less, to distinguish different levels of cognition. What's essential in the case of forms in the mind is the kind of information that they provide to the cognizing subject. So, to have distinct proper objects for different cognitive faculties, and also in order for us to be able to distinguish different levels of cognition, they have to operate on forms that represent different kinds of things.

The form at different levels of cognition must differ not only with respect to its modality (different modalities of forms at different levels of cognition are determined by how we are built), but also and most importantly with respect to its content. If so, then the FT in both its Numerical Identity and Formal Sameness versions cannot be correct. As I am going to show, this is exactly the conclusion that we are going to reach once we look more carefully at Aquinas's texts, especially those from De Veritate and from the Summa Theologiae. It is only the 'simplified Aquinas' that seems to require the FT. Aquinas's more advanced works suggest otherwise.
(v) Textual Evidence for Different Forms at Different Levels of Cognition

First of all, if we look closer at the passage from *ST* I 85, 1, ad1 (Q22 above), we can see that phantasms and intelligible species must differ in that there is something more in a given phantasm compared to the intelligible species that is abstracted from it. While the phantasm represents the object together with its individualizing conditions, or its 'individual principles,' there is nothing individualizing in the intelligible species. To say that there is something more in a phantasm compared to the corresponding intelligible species seems to suggest that they are actually not the same. If this is the case, we'd have to conclude that the FT model of cognition is not Aquinas's own.

We should recall that for Aquinas, individuating conditions of a thing are the material conditions of its form (cf. *InDA* III, 10, 706). Aquinas would say that intelligible species are completely separated from matter, and phantasms are still enmattered, that is, they are—to a greater or lesser degree—dependent on matter. A better way to clarify things would be to explain the 'individual principles' that phantasms represent as those features of a given thing that make it different from all other members of its species. Each individual cat that we encounter on the street is different from another. Each has fur of a slightly different shade or density, each has whiskers of a slightly different length; each has a different location, a different *hic et nunc*. These

---

58 *InDA* III 10, 706. "And the same is true of anything whose form exists in matter; there is something in it besides its specific principle. The specific nature is individualized through matter; hence the individualizing principles and individual accidents are not included in the essence as such. That is why there can be many individuals of the same specific nature—having this nature in common, whilst they differ in virtue of their individuating principles. Hence, in all such things, the thing and its essence are not quite identical. Socrates is not his humanity."
differences are registered in a phantasm (they don't have to be available to the consciousness of the cognizer, though), and they are omitted by the intellect (the concept CAT applies to every cat equally, no matter how much fur it has). In sum, phantasms represent individual features of individual members of a given species, and intelligible species represent those features that are common to the species as a whole. But if each vehicle of cognition represents different features of the object cognized, then they are certainly not the same, neither as mental items (they are not the same forms), nor as representations: what they represent, that is, their content, is different.

In addition, consider the following passages from Aquinas's *De Veritate* and from one of his two *Summas:

Q26. The name "intellect" derives from the fact that it has cognition of the intimate characteristics of a thing: for *intelligere* is by way of saying *intus legere* [to read penetratingly]. Senses and imagination have cognition of external accidents only; intellect alone succeeds in reaching a thing's essence. (*DV* 1, 12)

Q27. What is cognized by intellectual sight are the things themselves, not their images (*imaginæ*). This differs from bodily (sensitive) vision and spiritual (imaginative) vision. For the objects of imagination and sense are certain accidents from which the shape (*figura*) or image (*imago*) of the thing is made up. But the object of the intellect is the very essence of the thing, although the intellect cognizes the essence of the thing through its likeness, as through a means of cognizing, and not as through an object over which its vision is first carried. (*DV* 10, 4, ad1)) […] *obiectum intellectus est ipsa rei essentia; quamvis essentiam rei cognoscat per eius similitudinem, sicut per medium cognoscendi, non sicut per obiectum in quod primo feratur eius visio*

Q28. Sense is led through it [a form received from things] to a cognition of the external accidents; the intellect reaches to the bare quiddity of the thing, distinguishing it from all material conditions. Thus, when the mental cognizing is said to take its origin from sense, this does not mean that sense apprehends all that the mind cognizes, but that, from those things which sense apprehends, the mind is led on to something more. (*DV* 10, 6, ad2; cf. *SCG* IV 11, 15)

Q29. To understand is "to abstract the universal from the particular" (*ST* I 85, 1, obj.1). "The intelligible species is said to be abstracted from the phantasm; [but] not [in the sense] that the
identical form which previously was in the phantasm is subsequently in the passive intellect, as a body transferred from one place to another." (ST I 85, 1, ad3)

Based on the passages just quoted, and given the fact that the main function of cognitive faculties is to represent, we can now explicitly state the following claims concerning Aquinas’s account of cognition:

C4. Vehicles of cognition produced by the external and internal senses represent only accidental forms. (Q26 through Q28; "the objects of imagination and sense are certain accidents"; DV 10, 4, ad1).

C5. The intellect cognizes, or rather, represents only natures, or substantial forms, of things. As Aquinas puts it in another place, "the human soul knows the universal natures of things" (DV 10, 8) (Q26 through Q28). The intellect abstracts those essences from phantasms ("the intelligible species is said to be abstracted from the phantasm"; ST I 85, 1, ad3).

C6. The proper object of the intellect (the intelligible species) is abstracted from the phantasms (Q29).

Based on C4 through C6 we can infer:

C7. The intelligible species have content not present in sensory cognition59 (see again Q28: "from those things which sense apprehends, the mind is led on to something more" DV 10, 6, ad2).

The FT model would have to suggest that notwithstanding both C7 and C6, we must still hold:

C8. (as in Q29) The senses lead the intellect; they provide it with some determinate content.

The intellect's cognition is based on the information received by the senses in that the content of the intellect's proper objects is abstracted from phantasms, and so, in the end, it somehow comes from the senses.

59 This idea also in Geach, 1957, pp. 130–131.
The FT model's suggestion, however, has to be rejected. It just does not make sense to say that cognition consists in nothing but copying, and that, still, accidental features of things represented by the senses mysteriously appear as representations of essences at the level of the intellect. Rather, we should focus more on claims C4 and C5 which give an explicit answer to the worry mentioned above: they indicate that the external senses and the intellect do not operate on the same forms. On the contrary, they cognize very different things, the external senses cognize accidental forms of an object, and the intellect—its substantial or essential form. It is more plausible that, as we said above with respect to formal objects of different cognitive faculties, also here we should conclude that different things are represented at different levels of cognition.

The question that remains, of course, concerns the connection between the senses and the intellect. We'll need to figure out how sensible species and phantasms representing only accidental features of objects get transformed into intelligible species which supposedly represent the essences of things.

(vi) Abstraction does not work: The Transduction and the d/D Problems

As we said earlier, the usual way in which Aquinas's account of the activity of the intellect in its first operation is understood is as the process of abstraction: the intellect, supposedly, abstracts intelligible species from phantasms. Abstraction involves "disregarding the many accidental characteristics of a thing as preserved in the phantasm and focusing instead just on the thing's quiddity" (Stump, 1998, p. 292).
Consider, for instance, how in terms of abstraction we could describe the process of acquisition of the concept HORSE. Imagine that you never had any experiences with horses and that you do not have the concept HORSE. There are now four objects in front of you: A, B, C, and D. A is a thoroughbred (black, 17 hands high), B is an Arabian horse (bay coat, similar to the cow's color, 15 hands high), C is a Blonde d'Aquitaine cow (light brown), and D is a human baby. You observe the four things; you notice various individual features of each object, and also characteristics that are common to all or to some of them. All four things are alive, all of them move, make noises, and emit odors; all of them are mammals. They each have a different weight.

In order for you to arrive at the concept HORSE, you must ignore the fact that both things B and C are of the same color, that things A and C are of the same height, and that things A and D are both hungry. You must also ignore the fact that while A is 15hh high, B is 17hh, that while A has a long silky mane, B has its mane braided, or that while A is hungry, B is not. What you need to focus on is that both A and B are similar in important respects, both have similar shapes, both can run really fast, and both can carry a human on their back, etc.

This is exactly the kind of intellectual operation—in the above case, experiencing the various horse- and non-horse objects and attending to their horse-ness, while ignoring their other aspects—that is usually described as the process of abstraction.\(^{60}\) It is explained as a process of selectively attending to certain features and ignoring others.

---

\(^{60}\) Cf. Aquinas, *DSC* 10 where he explains: “one particular man, such as Socrates or Plato, makes things intelligible in act when he pleases, that is, by apprehending a universal form from particulars, when he separates that which is common to all individual men from those things which are peculiar to each.” Cf. also *ST* I 85, 1, ad1.
It is an act of 'isolated consideration' in which the intellect focuses on a feature (or a number of features) shared by a number of things and considers them apart from their points of difference.

The role of the intellect, on such an account, is to abstract the universal form HORSE by focusing only on some of the features shared by the thoroughbred and the Arabian horse, and ignoring all their other features. How plausible is it to claim, however, that a cognizer who didn't have the concept HORSE is able to focus on just the right features? How is his intellect supposed to know which features to choose unless it already had the concept HORSE? But if the previous possession of the concept HORSE is needed for the process of abstraction to work, then what we have here is a process of recognition of an object as a horse, and not a process of acquiring the new concept HORSE. The first problem with this traditional account of concept acquisition by means of abstraction is, therefore, that it is not clear that any new concepts can be acquired this way.

In addition, recall that for Aquinas the external senses provide us only with the information about the accidental features of a thing, and the intellect is supposed to operate only on things' essences. Here again it seems entirely implausible to suggest that the intellect manages to 'notice' a thing's essence just by looking at the information obtained from the senses.

There is no way to deny that Aquinas does talk about the process of concept acquisition as the process of abstraction (he says, for instance: "our intellect understands material things by abstracting from the phantasms," ST I 85, 1, and then he explains:
"This is what we mean by abstracting the universal from the particular, or the intelligible species from the phantasm; that is, by considering the nature of the species apart from its individual qualities represented by the phantasms," ST I 85, 1, ad1). And so I do not intend to deny this obvious fact. Still, I don't think that Aquinas is committed to the implausible views described above and implied by the traditional understanding of abstraction. We can keep the term abstraction—but we need to take into consideration the two problems mentioned above, and so, make some changes with respect to how the process of abstraction should be understood.

Abstraction as a process of concept acquisition cannot be considered an entirely empiricist process in which the intellect would arrive at a concept just by looking at the information provided by the external senses, ignoring some features, and focusing on others. On the contrary, there is a need for some innate mechanisms which will allow the intellect to produce a new concept when it considers the data provided by the senses.

1.2.4 Attempted (and failed) solutions

(i) No Platonic Nativism

Obviously, the transduction problem is a serious challenge for Aquinas's account. We do not know how we get intelligible species from perceptual data or how we get the essence of what a dog is from perceptual data given that the essence of a dog, according to Aquinas, is not specified by some description of perceivable properties like color or shape.
One position that Aquinas could have accepted in order to solve the transduction problem is some kind of Platonic nativism. This view would to some extent solve the problem of acquisition of concepts from sensible species. According to Plato, we have an innate knowledge of universal and eternal Forms, inhabiting the Platonic Heaven. These Forms correspond to concepts on which our intellect operates, or rather, to concepts that, on Plato's account, are themselves the object of the human higher cognition. The intellectual knowledge of the Forms is the only true and infallible knowledge. Plato believes that physical things, which are objects of perception, are in a constant state of change. They are only imperfect copies of the eternal and unchanging Forms. Our sensory knowledge, for Plato, is completely unreliable. It is not knowledge of the thing itself, but only knowledge of the copy. What's important is that for Plato sensory knowledge is not the source of intellectual knowledge. On the contrary, sensible species owe their identity to perfect and eternal Forms which they only resemble to some extent.

On Plato's nativism, there would be no problem with how to explain concept acquisition: concepts are not produced from sensible species; rather, they are innate in the human mind. The problem of transduction would also not be a threat for Plato because sensible species just don't get transformed into intelligible species. The senses are not where cognition starts. They actually constitute a hindrance against knowledge.

Aquinas, however, explicitly rejects Platonic nativism (cf., e.g., ST I 84, 1 and 4). He holds, as we have seen, that human knowledge takes its start from things, that is, it begins with the senses. The intellectual knowledge transcends the imperfect awareness of which other animals are also capable: "for to understand means to read what is inside
a thing. (…) the intellect alone penetrates to the interior and to the essence of a thing”
(DV 1, 12), but at the same time the intellectual knowledge is impossible without prior knowledge provided by the senses.

One reason why Aquinas rejects Platonic nativism is that it makes it impossible for us to know material things. If the only true knowledge was the knowledge of Forms, then science about the world that we inhabit would not be possible. In addition, Aquinas notices that it would be just ridiculous to introduce some extraordinary entities, utterly different from things that we experience in our everyday lives, in order to account for things that we do experience and which we perceive as pretty obvious.

For Aquinas it is clear that we need an account of cognition according to which physical things in the external world are what we know first. It is only through a reflective act that we can be aware of the cognitive form which is the product of our mind and the medium of cognition. Also, Aquinas is quite explicit in his conviction that the human intellect has a natural ability to arrive at intellectual knowledge of the actual physical things. The intellect has the power to abstract the universal from the individual.

The function of the human mind, on Aquinas's account, is to inform the cognizer about the external world. In order to ensure this, the mind needs to arrive at a correspondence with things in the world—it cannot create objects of cognitions, but it must assimilate itself to them (cf. 1.2.3 (ii)). (This on Aquinas's account is guaranteed by the fact that the senses are passive in their functioning, and that they are necessary for cognition.) If Plato was right, then "the senses would not be of themselves necessary for us to understand" (SCG II 76). Aquinas, however, is convinced that if God created
human beings with the external senses, this was for a reason. The reason for it was for us to be able to cognize the external world. If so, this is enough to prove that our senses are (for the most part) reliable, and that things in the world are such that it is possible for human beings to cognize them. The human cognizer was created as the composite of soul and body, and for this reason the body cannot be an obstacle to knowledge. On the contrary, the body is useful for knowledge, and it is also necessary: there would be no knowledge at all for human beings if they didn't have the senses.

(ii) Why not Illuminationism?

Another way in which one might try to solve the transduction problem while at the same time defending the view that the senses are a necessary step in the process of knowledge is to refer to an illuminationist view of cognition.

According to illuminationists, such as, for instance, Augustine, every normal human being is endowed with cognitive faculties including the senses and the intellect. Sensory cognition occurs in a natural way; our external senses are stimulated by information from the external world. The senses, however, only provide the raw data of cognition. Our human cognitive endowment on its own is not sufficient to lead us to certain and infallible knowledge of things. In order for higher cognition to take place, the intellect must be guided by divine illumination. Human concepts do take their origin in sensible species, but are not produced based on them. Rather, they are copies of
divine ideas, or 'un-created exemplars.' (Divine ideas are templates for human concepts.) In human cognition, there is still no natural passage from sensible species to concepts. For John to be able to arrive at the concept CAT when he has sensory experiences of cats, the sensible species produced by his external senses has to be divinely illuminated. As a result of this illumination, the given sensible species is matched up with the corresponding exemplar. The outcome of this process is the concept CAT in John's head. Without divine help John's cognition would be limited to what's available to animals, i.e., to sensible qualities of things.

Even though on illuminationist account sensory experience is necessary for knowledge acquisition, human knowledge and concepts still do not really come from experience. Concepts are in us innately, or they reside somewhere in God's mind. In any case, they are not acquired in a natural way. They need to be revealed to us by some kind of divine illumination. For illuminationists, therefore, the transduction problem is solved at a supernatural level. The intellect does not need to perform any mysterious operation. The job is done for it by the divine mind.

The idea that Aquinas might adopt illuminationism at least at first seems very surprising given that Augustinian illuminationism is very similar to Platonic nativism (the difference between them seems to be merely that illuminationism replaces the Platonic Forms with the divine exemplars; cf. ST I 84, 5). Since Aquinas so explicitly rejected Platonic nativism, it seems that he would be equally critical of illuminationism.

---

61 I'll save the term 'concept' for the ideas in the human mind; divine ideas will be referred to as exemplars.
62 It would also be limited to mere opinion: it would not reach truth or certainty.
But this, according, for instance, to Pasnau (2002), is not so obvious. There are passages in Aquinas's own writings which suggest an illuminationist interpretation. Aquinas talks about the "light of the agent intellect" and compares it to "a participating likeness in the uncreated light, in which all the divine ideas are contained" (cf. *ST* I 84, 6). Also, in *De Veritate* Aquinas explains:

Q30. Our mind receives knowledge from sensible things; nevertheless, the soul itself forms in itself likenesses of things [*ipsa anima in se similitudines rerum format*], inasmuch as through the light of the agent intellect the forms abstracted from sensible things are made actually intelligible so that they may be received in the possible intellect. (*DV* 10, 6)

In the quoted passage Aquinas seems to be suggesting that even though the senses do constitute the first step in the process of cognition, still the cognizer's soul, or his mind, produces its own vehicles of cognition independently of the senses ("the soul itself forms in itself likenesses of things"). This happens with the help of what Aquinas calls the "light of the agent intellect" through which "the forms abstracted from sensible things are made actually intelligible." The agent intellect is some kind of light that transforms phantasms into intelligible species.

According to Pasnau, the passage from *DV* 10, 6 (Q30) indicates that in Aquinas's views empiricism coexists and collaborates with an Augustinian doctrine of illumination. First of all, Pasnau notices, "Aquinas's Augustine remains thoroughly distinct from Aquinas's Plato" (Pasnau, 2002, p. 304). For Aquinas's Augustine, divine ideas are not objects of cognition, but only its sources, and the senses play a crucial role
in cognition. This version of Augustinian illuminationism, therefore, seems much more plausible than the views that Aquinas ascribes to Plato. Aquinas also says:

Q31. all knowledge is in a certain sense implanted in us from the beginning (since we have the light of the agent intellect) through the medium of universal conceptions which are immediately known by the light of the agent intellect. These serve as universal principles through which we judge about other things, and in which we foreknow these others. (DV 10, 6)

According to Pasnau, this passage shows that

Q32. Without appealing to God, Aquinas sees no way of explaining how we acquire our most basic concepts and recognize the truth of first principles. Neither deductive nor inductive reasoning can account for how we immediately see that such principles are true. This insight is simply something we are given. (Pasnau, 2002, p. 309)

If Pasnau is right, then in the above passage from DV 10, 6 Aquinas is explicitly denying both empiricism and naturalism of cognition. Aquinas is thus not a naturalist about the way in which the mind works. Rather, according to Pasnau, Aquinas's goal is "nothing less than to affirm Augustine's theory [of divine illumination] and place it at the very heart of his own account of intellective cognition" (Pasnau, 2002, p. 307).

There are, nonetheless, serious reasons why we might want to question Pasnau's conclusion. Pasnau himself realizes at some point that what Aquinas seems to be positing in his account of cognition are certain innate capacities or innate mechanisms: for Aquinas, "we must possess the innate capacity to see the truth of certain principles. If our mind were entirely blank, our education could never begin" (Pasnau, 2002, p. 308).

---

63 Pasnau also contrasts Aquinas's account of cognition with Duns Scotus's naturalism (Pasnau, 2002, p. 310) and explains that the difference between the two philosophers is that for Scotus it is the world that is intelligible; "the world is illuminated" and this is sufficient for cognition. For Aquinas, on the other hand, it is the human cognizer that needs to be illuminated in order for cognition to be possible.
But then, the transition that Pasnau makes from these possible 'innate capacities' to an illuminationist view that denies naturalism is far from obvious. Rather, it seems more plausible (and this is an interpretation that Pasnau also considers), that Aquinas has only a specific kind of illumination in mind—one that has occurred

Q33. at the time when the soul was created. (…) God does its part at the outset furnishing human beings with a sufficient capacity for thought on their own, without the need for "new illumination added onto their natural illumination" (ST I-II 109, 1). (Pasnau, 2002, p. 305)

This should be contrasted with what Pasnau calls "special illumination, an ongoing influence from above, constantly required for the intellect's operation" (Pasnau, 2002, p. 305). Aquinas, therefore, may be simply saying that God created human beings with some special innate cognitive capacities. Innateness, however, does not exclude naturalism.

In addition, when Aquinas directly considers the question of whether the human mind needs divine illumination in order to know the truth, his answer to it is, decidedly, 'no' (cf. quotes from InBoeth 1, 1 below). Several reasons might be identified that would justify such an answer.

A. First of all, the illuminationist explanation of cognition is just not needed. Aquinas holds firmly "that the human mind can see the truth by its natural light without anything being added" (InBoeth 1, 1). He admits that "the human mind is divinely illumined by a natural light" but, he emphasizes, this kind of illumination is very different from that which is posited by genuine illuminationists. We can indeed call our human intellect a
light or an illumination. This is because it allows us to see so much more than what is accessible to animals—it allows us to understand and to reach the essences of things. But what all this points to is really the intellect's natural power. For Aquinas it is obvious that we are somehow illumined by God in everything we do, and so also in our knowledge. This is because we are created by him and depend on him in everything. But God created us with our natural power of intellect able to reach the truth about the world. So, Aquinas explains, "that light whereby our mind understands (...), the agent intellect (...), is something of the nature of the soul" (DSC 10, otc3).

B. Secondly, the illuminationist account of cognition leads to a regress. Consider again the passage from Aquinas's commentary on Boethius's De Trinitate:

Q34. On the contrary, the human mind is divinely illumined by a natural light, according to Psalm 4:7: "The light of your countenance, O Lord, is signed upon us." Therefore, if this light, because it is created, is not adequate to know the truth, but needs a new illumination, the added light with equal reason will not suffice, but will require another light, and so on to infinity—a process that can never be completed. And so it will be impossible to know any truth. Therefore we must depend on the first light, so that the human mind can see the truth by its natural light without anything being added. (InBoeth 1, 1, otc1)

If we do not have a natural power to know the truth, then it seems that knowledge would just not be possible.

In addition, negation of naturalism in cognition, by requiring the constant intervention of the Creator, "derogates from the order of the universe, which is made up of the order and connection of the [natural] causes" (DV 11, 1). For Aquinas, God created the universe in such a way that natural causes are sufficient for cognition.
C. Finally, the illuminationist explanation seems useless because it does not provide any positive account of cognition. Illuminationism basically says that whenever we conceive the concept CAT, something like the Holy Spirit must have intervened in order to make sure that the concept that appears in our mind matches the cat in the world that we have experienced. But this just amounts to repeating that (i) since we do in fact think about things in the world, there must indeed be a link between our thoughts and things in the world, and that (ii) we just do not know how it works.

Aquinas would not use illuminationism in order to solve the transduction problem or the d/D problem. He explains that

Q35. our mind has within it the power to make objects actually intelligible, namely the agent intellect, and an intelligible object of this sort is proportionate to the intellect. So the mind does not need a new illumination in order to know the truth. (InBoeth 1, 1, otc2)

Q36. Intellectual light is related to mental sight [i.e., understanding] as physical light is to bodily sight. But every physical light, no matter how weak, makes something be seen physically—at least itself. Therefore the intellectual light that is connatural to the mind also suffices to know some truth. (InBoeth 1, 1, otc3)

Q37. Just as the other natural active powers, joined to their passive counterparts suffice for natural activities, so also the soul, endowed with an active and passive power, is adequate for the perception of truth. (InBoeth 1, 1)

For Aquinas, the capacity to acquire concepts in response to experiences with things in the world is to some extent similar to, say, talent for music. Surely, Aquinas would agree that musical talent is God-given, as is everything else in the world. He says in In Boeth. De Trinitate 1, 1: "all created active powers function under the movement and direction of the creator." But when we want to explain what it is that makes young Mozart play beautifully whenever he sits at the piano, we do not need to have recourse to the help of
the Holy Spirit. Musical talent is Mozart's innate endowment. Similarly, the 'direction' of
the Creator with respect to human cognition does not happen on a daily basis. No
additional supernatural help is needed on Aquinas's account when we want to explain
what it is in virtue of which John's experience with cats leads him to acquire the concept
CAT, and also what it is that guarantees that when John's senses register something
furry, and gray, and moving, he will entertain CAT- and not BICYCLE-thoughts, or that
when he conceives of the concept CAT, John's thoughts are about cats and not about
butter. Rather, what I want to suggest is that this is done through some instinct-like innate
mechanisms.

Therefore, my claim is that we should translate the mysterious light of the agent
intellect (as in DV 10, 6; QQ30–31 above) into innate endowment of our cognitive
system. Instead of saying that cognition is rendered possible by some divine
illumination, we will say that there are in the cognizing agent certain cognitive
mechanisms fixed in the species by natural selection, and inherited by all the members
of the species. The agent intellect collaborating with the cogitative power is this kind of
innate mechanism functioning according to a set of innate rules. (All this will be
explained in more detail below. For a more detailed explanation of innateness that I
propose, see 5.2.4 and 5.2.5.)
CHAPTER II
INNATE ELEMENTS IN AQUNAS: THE LEVEL OF THE SENSES

2.1 The need for innate elements in Aquinas's account of cognition

Recall the passages from Aquinas's *De Veritate* and from *Summa Theologiae* quoted above in 1.2.3 (v):

Q26. The name "intellect" derives from the fact that it has cognition of the intimate characteristics of a thing: for *intelligere* is by way of saying *intus legere* [to read penetratingly]. Senses and imagination have cognition of external accidents only; intellect alone succeeds in reaching a thing's essence. (*DV* 1, 12)

Q27. What is cognized by intellectual sight are the things themselves, not their images (*imaginæ*). This differs from bodily (sensitive) vision and spiritual (imaginative) vision. For the objects of imagination and sense are certain accidents from which the shape (*figura*) or image (*imago*) of the thing is made up. But the object of the intellect is the very essence of the thing, although the intellect cognizes the essence of the thing through its likeness, as through a means of cognizing, and not as through an object over which its vision is first carried. (*DV* 10, 4, ad1) [... *obiectum intellectus est ipsa rei essentia; quamvis essentiam rei cognoscatur per eiu similitudinem, sicut per medium cognoscendi, non sicut per obiectum in quod primo feratur eius visio*]

Q28. Sense is led through it [a form received from things] to a cognition of the external accidents; the intellect reaches to the bare quiddity of the thing, distinguishing it from all material conditions. Thus, when the mental cognizing is said to take its origin from sense, this does not mean that sense apprehends all that the mind cognizes, but that, from those things which sense apprehends, the mind is led on to something more. (*DV* 10, 6, ad2; cf. *SCG* 4, 11, 15)

Q29. To understand is "to abstract the universal from the particular" (*ST* I 85, 1, obj.1). "The intelligible species is said to be abstracted from the phantasm; [but] not [in the sense] that the
identical form which previously was in the phantasm is subsequently in the passive intellect, as a body transferred from one place to another." (ST I 85, 1, ad3)

We said that these passages, together with the fact that for Aquinas the main function of cognitive faculties is to represent, imply the following claims:

C4. Vehicles of cognition produced by the external and internal senses cognize only accidental forms. (Q26 through Q28; "the objects of imagination and sense are certain accidents"; DV 10, 4, ad1).

C5. The intellect cognizes, or rather, represents, only natures, or substantial forms, of things. As Aquinas puts it in another place, "the human soul knows the universal natures of things" (DV 10, 8) (Q26 through Q28). The intellect abstracts those essences from phantasms ("the intelligible species is said to be abstracted from the phantasm"; ST I 85, 1, ad3).

C6. The proper object of the intellect, the intelligible species, is abstracted from the phantasms (Q29).

And that C4 through C6 imply:

C7. The intelligible species have content not present in sensory cognition (as in Q28: "from those things which sense apprehends, the mind is led on to something more" DV 10, 6, ad2).

The FT, as we said, would also claim

C8. The senses lead the intellect and provide it with its proper object.

We already pointed out above that the interpretation suggested by the FT should be rejected. The FT holds that we need to support the claim C8 if we want to preserve Aquinas's cognitive empiricism (CE): we need to say that the intellect is guided by the senses and that what's in the intellect is the same as what was in the senses. Aquinas indeed believes that the senses must supply some cognitive content to the intellect for any cognition to take place. He does not need to claim, however, that the whole content of vehicles of cognition at the intellectual level is provided by the senses.
The passage from DV 10, 6, ad2 (Q28) indicates that the FT interpretation is wrong because there is an important difference between the kinds of forms operated on by sensory and intellectual cognitive faculties. As it was discussed above (cf. 1.2.3 (iv)), the formal object of the intellect, on Aquinas’s account, is intelligible species, that is, a concept, a cognitive form that represents quiddities, or essences of things. On the other hand, the formal object of the external senses is the cognitive form representing accidental features of things. And so, as is clearly seen in the claims C4 and C5, the external senses and the intellect do not operate on the same forms. On the contrary, they cognize very different things, the external senses cognize accidental forms of an object, and the intellect—its substantial or essential form.

Indeed, if we consider together the claims C4 and C5, we notice that there seems to be no connection between the level of the senses and the level of the intellect. It is a mystery how the intellect could abstract universal essences based on just the accidental features of things or how it is possible to go from sensory stimuli to concepts expressing essences. Vehicles of cognition from the sensory level of cognition seem completely different from concepts: there is more, and a different kind of information in concepts than there was in sensible species and phantasms. The problem of transduction, therefore, looks even more serious now than it seemed before.

Given that Aquinas believes that the intellect cognizes the essences of things, it is not clear how the intellect’s cognition is still supposed to be somehow based on what’s provided by the senses. It does not seem plausible that sensible species representing accidental features of the object cognized would suddenly turn into an intelligible
species representing the object's essence. Similarly, it is not clear how the intellect could
derive its knowledge of essences from the information provided by the senses. Consider
the following case. A part of the essence of, say, a wolf is that it is a non-rational animal.
This essential feature of wolf-ness would be cognized by the intellect. The intellect is
supposed to arrive at this object of cognition on the basis of the data provided by the
external senses. Non-rationality, however, is quite different from colored expanses and
discrete tastes, and any other quality that can stimulate the external senses. It does not
seem plausible to say that non-rationality is just another element among other sensible
features. So how can it be that the intellect cognizes non-rationality of wolves based on
the information provided by the senses? In addition, given that on the first two versions
of the FT account the form is the same at every level of cognition, the external senses
themselves would already have to cognize the wolf's essence. Otherwise we would have
to deny the possibility of ever cognizing such features as non-rationality.64

The solution that Aquinas could offer to the problem of transduction is indicated,
I believe, by the claim C7. C7 says that there is something in the intellect that was not
previously in the senses, and so it contradicts the FT account. What it also indicates is
the need to interpret Aquinas's account of cognition as involving certain innate elements.
Since the additional content of the intellect does not come from the senses, it has to be
either innate, or acquired in some innate way.

64 See below, 3.2, for a more detailed analysis of what it means, for Aquinas, to cognize essences of things. See also 3.1 and 3.3 where I offer a solution to what I called the Concepts as Essences Problem (announced in 1.1.3 (iv)).
The rejection of the FT, as we said above, does not threaten Aquinas's empiricism. Still, it would certainly not be plausible to claim that Aquinas is at the same time an empiricist and a nativist with respect to the same aspects of cognition. Aquinas would not give up his empiricism. This is the view that he took from Aristotle, a view that ensures Aquinas's naturalism and which places human beings within the realm of animals, of embodied living organisms endowed with sensation. This is also a view that gives enough emphasis to the importance of our senses in cognition, and which indicates that sensory knowledge is a necessary source of our higher level cognition. Empiricism, finally, helps to guarantee the objectivity of cognition: it guarantees that there is a link between concepts in the cognizer's mind and things in the external reality. The nativist element in Aquinas's account of cognition, on the other hand, negates the first two versions of the FT interpretation. This way it is able to provide grounds for distinguishing different cognitive faculties in the first place. But if such nativism is indeed present in Aquinas's views, and if all three kinds of cognitive faculties operate on different forms, then it is not clear that the continuity of the process of cognition is preserved or what exactly guarantees it: how, for instance, is the cat (or catness) cognized by the intellect connected to whatever of the same cat is cognized by the internal and the external senses?

And so, there are two elements that need to be reconciled in Aquinas. First of all, it is crucial that the objectivity of cognition is preserved. Secondly, because of the differences between concepts of our intellect and the vehicles of cognition from the level of the senses, we will have to figure out a way to guarantee that our intellectual
cognition still has something to do with the world cognized by the senses, without the help from Platonic nativism or illuminationism.

The solution to these problems is to be found in the Form Trans-Formation model which emphasizes the role of the internal senses in Aquinas's account of cognition. Let us immediately admit that it may not be obvious that introducing internal senses is helpful here instead of only adding to the confusion. On Aquinas's account, the proper object of the internal senses is supposed to be different from both the proper object of the intellect and of the external senses. In that case, we are not better off but worse: the problem has doubled. There seemed to be no explanation of how accidental forms of the external senses can turn into substantial forms of the intellect. Now it turns out that our ignorance is even wider: it concerns both (a) the relationship between sensible species and phantasms (we don't know how some discrete units of sensory data registered by sensory organs turn into images of individual substances), and (b) the relationship between phantasms and intelligible species (we don't know how images of individual things turn into universal concepts).

I suggested that the claim C7 is an expression of Aquinas's nativism to the effect that the additional content of intelligible species must be innate, or that there must be an innate mechanism that produces this additional content. The mystery about the relationship between sensible species and phantasms indicates, I believe, that not only is there something in the intellect that was not previously in the internal senses, but also there is something in the internal senses that was not previously in the external senses. So it turns out that on Aquinas's account there is something innate not only at the level
of the intellect but also at the level of the internal senses. In fact, Aquinas himself hints at this very idea in the *Summa Theologiae* when he says:

Q38. Although the operation of the intellect has its origin in the senses: yet, in the thing apprehended through the senses, the intellect knows many things which the senses cannot perceive. In like manner does the estimative power, though in a less perfect manner. (*ST* I 78, 4, ad4)

Therefore, what first seemed to be a doubled problem is at the same time an indication of a solution. The mystery from the level of the internal senses can be solved once we recall that for Aquinas the internal senses are like instinct. Instinct is something innate, and similarly the internal senses can be understood as innate cognitive mechanisms. This explanation will then be also applied to the intellect. We are going to suggest that the intellect as well is an innate mechanism. It is an evolutionary adaptation for the human beings that they are capable to create concepts and to assign them to phantasms produced by the higher internal senses.

2.2 The Form Trans-Formation interpretation—A first approximation

As we already indicated, the interpretation that I propose, that is, the Form Trans-Formation interpretation of cognition, can also be called the *Similarity* version of the FT.

---

65 The parallel between the relationship between sensible species and phantasms, and the relationship between phantasms and intelligible species confirms my intuition that what we want in order to solve the transduction problem in Aquinas's account of cognition are innate mechanisms, and not innate concepts. We don't know how phantasms are acquired from sensible species. But we don't want to solve this problem by suggesting that phantasms are innate. Similarly, the difficulty with explaining how intelligible species are acquired from phantasms is not to be solved by suggesting that intelligible species are innate. In both cases we rather want to posit certain innate mechanisms.
On this account, there is neither numerical, nor formal identity between the forms involved in the cognitive process.

(i) The Process of Cognition According to the FTF

The FTF does not differ from the FT with respect to how it describes the first stage of the process of cognition. According to the FTF account, in the first stage of the process of cognition the external senses are affected by the species impressae of sensation, that is, by some kind of configuration of photons of light, sound waves, odor molecules, etc. The species expressae of sensation produced by the external senses represent (and send to the brain) the information: 'black,' 'moving,' 'warm,' 'meowing,' 'furry,' etc. On this basis, the internal senses produce a phantasm representing clusters of proper and common sensibles, or, in other words, a unified percept representing the accidental features of one and the same object. This is where the FTF starts to differ from the FT. Already at the level of the internal senses the form that is received from the external senses has to be transformed into a proper phantasm; it is not merely a copy (or a simple sum) of forms operated on by the external senses. When the cogitative power takes in the phantasm representing clusters of sensible qualities, it can produce a phantasm that represents the substance that these qualities are accidents of, and that represents the substance as an individual thing of a kind, only because the cogitative power acts according to its innate

---

As we will see below in more detail, for Aquinas, the role of the internal sense of cogitative power is to apprehend the individual thing 'as an existing thing of a common nature' (InDA II 13, 398), and to be "aware of a man as this man, and this tree as this tree" (InDA II 13, 398). Unlike all other sensory powers, the cogitative power does not cognize accidental features of things. Instead, its proper objects are incidental sensibles, and specifically – substances.
programming. To this substance the cogitative power applies one of the intellect's concepts, or intelligible species.

The cogitative power is a *sensory* faculty and what particular substance it detects is determined by the input it receives from the common sense. *How* it apprehends the given substance, on the other hand, is determined by the intellect. The phantasm that the cogitative power receives from the lower internal senses represents the unified percept encoding the information: 'at the same time black, moving, warm, meowing, furry, etc.' The phantasm that the cogitative power produces is already influenced by the intellect (see 3.3 below). It encodes the information: 'this here is an animate thing that is cat-like black, moving, warm, meowing, furry, etc.' (Or perhaps we could say, using Chisholm's terminology, that it encodes the information: 'this here is an animate thing that is apprehended cat-ly.')

(ii) A-symmetrical Relation: Causal Connection and Representation

On the FTF account, the cognitive form that appears in the intellect (Fi), the form CAT, is different from the original form informing the object of cognition (Fo), from the sensible species in the sense organs (Fs), and from the phantasm in the (lower) internal senses (Fis).^67

The relationship between forms involved in cognition goes as follows: the form in the external senses is obtained from the form of the object by means of some transformation mechanisms (Fs = TR(Fo)). The form in the lower internal senses (the

---

^67 Cf. 3.1.1 (ii) for an account of different kinds of phantasms that should be distinguished.
phantasm) is obtained from the sensible species (Fis = TR(Fs)), and the form in the intellect (Fi) from the form in the internal senses (Fi = TR(Fis)). Fo gets transformed into Fs, Fs into Fis, and the resulting form, Fis, gets transformed into Fi; therefore, Fo ≠ Fs ≠ Fis ≠ Fi. What we have here, however, is both cognitive empiricism and nativism. The empiricism of the position consists in the fact that Fi is defined in terms of Fis, Fis in terms of Fs, and Fs in terms of Fo. (It is what we could call a 'deeper kind of cognitive empiricism,' that is, the kind of empiricism which says that what is in the intellect was not just in the external senses, but rather in the external objects themselves.) The nativism proposed by the FTF does not concern innate concepts but rather posits two levels of innate mechanisms of transformation of forms. We still call these forms 'similar' because of the fact that there is something common at every level of cognition, and that there is a special causal link between all the forms involved, with the causal process going just one way.

The intellectual form (Fi) is neither literally nor formally identical to any of these other cognitive forms. Aquinas explains:

Q39. One kind of assimilation occurs because of an agreement in nature (…) but this kind is not required for knowledge. Another kind of assimilation occurs through information, which kind is required for cognition—just as sight is assimilated to color, when the pupil is informed by its species. [dicendum, quod est quaedam assimilatio secundum convenientiam in natura; (…) sed haec non est illa quae requiritur ad scientiam. Est etiam quaedam assimilatio per informationem, quae requiritur ad cognitionem; sicut visus assimilatur colori, cujus specie informatur pupilla.] (In I Sent 34, 3, 1, ad4; cf. also DV 2, 3, ad9)

Q40. That application of the known to the knower, which causes knowledge, should not be understood by way of identity but rather by way of representation. Therefore, it is not necessary that the mode of the knower and of what is known be the same (DV 2, 5, ad7)
For Aquinas, forms can be related to each other either 'through an agreement in nature' or 'through information.' In the latter case, one form can be a likeness of (or be similar to) another just in case it represents it. As Aquinas explains, "to represent something [just] is to be the likeness of that thing" (DV 7, 5, ad2: *repraesentare aliquid est similitudinem eius continere*). And so, we could say that the relationship between the form in the mind and the form of the object consists exactly in the former *representing* or *signifying* the latter.

We call the two forms that are related to each other in this way 'similar' not just because that's how Aquinas talks about it. The *similarity* of these forms has a specific meaning: it does not refer to numerical or formal identity, but rather to a causal connection between the two forms, and specifically to a causal relation between the object cognized and the species in the mind of the cognizer.68 Recall the following passage from the *Summa Theologiae*:

> Q41. If, however, the agent and its effect are not contained in the same species, there will be a likeness, but not according to the formality of the same species; as things generated by the sun's heat may be in some sort spoken of as like the sun, not as though they received the form of the sun in its specific likeness, but in its generic likeness. (*ST* I 4, 3)

---

68 This would be an instance of the so-called equivocal causation. As Kretzmann and Stump explain (Kretzmann and Stump, p. 334), for Aquinas efficient (non-accidental) causation always involves an agent (A), a patient (P), and a form (f). The efficient cause (A) *somehow* possesses the form (f), and then, "A's exercising causal power on P brings about f in P, somehow." The efficient cause (A) and the effect (P) can have the form (f) in several different ways (this is what is brought out by the term 'somehow'). In straightforward (or *univocal*) efficient causation, "A and P have f in just the same way (…); f can therefore be predicated truly of each in just the same sense. The metal hotplate and the metal kettle bottom resting on it are both called hot univocally: the form heat in these two causally related objects is the same specifically and differs only numerically." In *equivocal* causation, "there is no obvious respect in which to say that the f effected in P is found antecedently in A, and yet there is a natural causal connection (…). If (A) is solar power and its effect is the hardening (f) of some clay (P), then obviously the sun's power is not itself hard, as the clay is. To say what it is about solar power that hardens clay will not be as easy as explaining the heating of the kettle, and yet the hardening of the clay must, somehow, be brought about by that power. In such a case, A has f only in the sense that A has the power to bring about f in P."
When two things are different, but still related to each other in that one is the cause of the other, this is enough to call them similar. This is because on Aquinas's account,

Q42. an effect manifests in some way its cause, since it proceeds from the cause according to an intelligible pattern. A cause communicates something of its perfection to the effect, expresses itself in the effect, which therefore must be to some degree similar to its cause. (Velde, 1995, p. 93)

That is why it seems plausible to still talk of similarity in the case of an object in the external world causing a mental representation in the cognizer's mind. What makes them 'similar' is the causal link between the form informing the object in the external world (Fo) and the form in the mind (Fi). We call them 'similar' also in virtue of the fact that the form in the mind (Fi) is a representation of the object (Fo). (The two forms are similar perhaps in the same way in which we would say that the rings on the tree trunk are 'similar' to the age of the tree. There is certainly neither numerical nor formal identity between the two items. The age of the tree is the cause of the rings on the tree trunk, but the age of the tree does not represent the rings. The rings represent the age of the tree, but do not cause it.)

On the first two versions of the FT account the content of the form at different stages of the process of cognition doesn't change; whatever gets scanned by the sense organs is also copied to the higher level cognitive faculties. On the FTF, or the Similarity version, on the other hand, the intellect's work is still done on the basis of the sensible forms, but it is creative: its end product, that is, concepts, contain something that was not present in phantasms produced by the lower internal senses. In a similar way, there
is neither numerical identity nor sameness between sensible species and phantasms. Those phantasms that the cogitative power operates on contain different information than the sensible species that give rise to them.

(iii) Innate First Concepts vs. Innate Mechanisms of Trans-Formation

It seems obvious that the main challenge for the FTF is to come up with a plausible account of the form trans-forming mechanisms that would allow us to solve the transduction problem. Someone might argue, however, that what Aquinas needs, and what can be actually found in his texts is the presence of some innate concepts. Before we go on presenting the idea of innate cognitive mechanisms in Aquinas's account, let us see why the idea of innate concepts should be rejected.

Recall the passage from De Veritate, where Aquinas says that we acquire knowledge

Q43. through the medium of universal conceptions which are immediately cognized by the light of the agent intellect. (DV 10, 6)

One way to interpret this passage is that acquisition of knowledge is possible because there are in the mind some innate universal conceptions which do not need to be acquired, but are already there. These 'universal conceptions,' Aquinas continues,

Q44. serve as universal principles through which we judge about other things, and in which we precognize these others. In this respect, that opinion is true which holds that we previously had in our cognition those things which we learn. (DV 10, 6)
We said above that it would be incorrect to hold an illuminationist interpretation of Aquinas's account of cognition or to ascribe to him Platonic nativism. However, even if we agree that those two explanations are to be rejected, there still seem to be two ways in which the above passage could be interpreted.

A. First of all, one could claim that the universal conceptions immediately cognized by the intellect are *innate first concepts*. We discussed above Aquinas's explicit rejection of Platonic nativism. One could still argue, however, that rejecting Platonism is only rejecting innate concepts of *natural kinds*, and that although for Aquinas concepts of natural kinds are indeed *not* innate but rather acquired from the phantasms through abstraction, there are also for him some more fundamental concepts, some first conceptions, which are innate.

The 'universal conceptions' that are immediately available to the intellect would be interpreted as innate first concepts, such as, for instance, the concept IDENTITY, BEING, ONE, GOOD, etc. These concepts are constitutive elements of *first principles*, that is, of some self-evident statements which express the fundamental laws governing the functioning of the human intellect. If the first concepts are available immediately, that is, without the help of experience, then the intellect must have them innately. Those innate first concepts would be considered a necessary condition for cognition.

B. Secondly, and this is the path that I'm going to follow, the quoted passage could be interpreted as suggesting that the immediate cognition through the light of the agent
intellect is really an *innate mechanism*, a programmed reaction of the mind to the data provided by the senses.

I am going to claim that:

C9. Less innate stuff is better than more.
C10. Innate concepts would still require innate mechanisms, but not vice versa.
C11. Aquinas's account does require innate cognitive *mechanisms*. And in fact, innate cognitive mechanisms are already implicit in Aquinas. There is no need to posit in addition innate concepts.

I am going to assume C9 and C10 without providing much support for either. But just briefly, we should accept C9 because positing innate stuff is, I believe, a sign of capitulation. It is when we can find no way to explain how something was acquired from experience that we admit that it must be innate. So, the more innate stuff we posit, the more we must have failed.

As for C10, recall our discussion of illuminationism and Platonic nativism. In both cases knowledge and concepts do not come from experience; we have them *independently* of experience (experience may only be needed to activate them). But they are still supposed to represent things that we encounter *in* experience. If experience does not provide a link between a given concept and the object it represents, then there are two choices: either we try to pull something out of a hat (e.g., divine intervention), or we must still posit some innate mechanisms to match the two, and to solve in this way the
transduction problem. Innate concepts alone are not sufficient for a coherent theory of cognition. There seems to be a need to posit innate mechanisms no matter what.

The literal interpretation of ‘universal conceptions immediately cognized by the intellect’ is to be rejected. Even if we agreed that there are some kinds of innate first concepts on Aquinas’s account, we’d still not know how concepts of natural kinds are acquired. And what we are mostly concerned with is exactly how Aquinas could explain that concepts of natural kinds are acquired from sensible species. Secondly, on the interpretation of concepts as thought-parts, and this is the interpretation that I propose (see 3.1), innate first concepts just do not make sense. BEING, ONE, GOOD, etc., are certainly not what human beings think first. On the contrary, as we are going to see, many people never even entertain thoughts containing these concepts.

And so, what’s crucial for my considerations—and what we should keep from the whole discussion of the FTF account of cognition—is the need for a specific kind of a causal link between the object and the form in the mind. What we need is an asymmetrical relationship between forms involved in the process of cognition consisting of formal causation going one way, and representation going the other way. This is exactly what the FTF emphasizes. Formal causation in Aquinas’s account of cognition does go only one way, from object to mind, because (in typical cases) cognition starts with sensation, and because external senses are passive. In addition, the idea of transformation of forms indicates that it is really the object together with the architecture of the mind that determines the informational content of our concepts, contrary to the

\[69\] This of course will in the end depend on what account of concepts one endorses. Cf. below, 3.1.
two versions of the FT, which would suggest that it is just the object that is supposed to do that.

The claim that I want to make in this dissertation is that when Aquinas talks about *innate conceptions* or *innate first concepts*, what he means by it is that no cognition is possible without transcendentals which should be understood as innate mechanisms that determine the functioning of the intellect and that require the collaboration of the intellect with the cogitative power.

Before I provide a more detailed presentation of the FTF's interpretation of the process of cognition, several preparatory steps will be needed. First, we are going to discuss Aquinas's view concerning the internal senses. It is at this level that Aquinas himself seems to indicate that the human cognition functions by means of certain innate mechanisms. Then we'll turn to what I consider the second level of innateness in which the internal sense of the cogitative power collaborates with the intellect by cognizing the world through the intellect's transcendentals, and by applying the intellect's concepts to the phantasms that it produces.
2.3 The role of the internal senses and instinct

2.3.1 Internal senses in general: Terminological confusion

In the *Summa Theologiae* I 78, 4, Aquinas describes four internal senses:

1) *sensus communis* is directed, like the external senses, towards receiving sensible forms;
2) *phantasia* or imagination (*vis imaginativa*) is "directed at the retention or preservation of sensible forms"; it "serves as a kind of treasury for forms grasped through the external senses" (cf. also *DV* 10, 2);
3) the *estimative/cogitative power* is "directed at apprehending intentions that are not grasped through the external senses," that is, *intentiones non-sensatae*\(^{70}\), such as danger, and at forming instinctive judgments, e.g. that an animal should flee:
   a. in animals that do not have reason, it is called the estimative power, and is usually understood as some kind of an instinct;
   b. in rational beings, it has the name of the *vis cogitativa*;
4) *sense memory* (*vis memorativa*) serves as a treasury for *intentiones non-sensatae* that are apprehended by the estimative power (*ST* I 78, 4).

---

\(^{70}\) *Intentiones non-sensatae* is another term for incidental sensibles, i.e., intentions which are not perceived through the senses (*ST* I 78, 4). The term 'intention' comes from a distinction proposed by Avicenna. According to Avicenna, there are two kinds of sensory objects, forms and intentions: "A form is that which (...) an external sense first apprehends, and then gives to an internal sense—for example, when a sheep apprehends the form of a wolf: its shape, condition, and color (...). An intention is that which the soul apprehends of the sensible, even though the external sense has not previously apprehended it—for example, when a sheep apprehends the intention that it has of the wolf: that it ought to fear it and flee from it—even though its senses do not in any way apprehend this" (Avicenna, *Liber de Anima* I 5, p. 86 as quoted by Pasnau, 2002, p. 269). In Aquinas, however, we see the term 'form' used interchangeably with the term 'intention.' Therefore, we can say that while *intentiones non-sensatae* stand for the proper objects of the internal senses (they are forms that cannot be directly cognized by the external senses), the external senses operate on *intentiones sensatae*, i.e. on sensible species, i.e., on proper or common sensible forms.
Aquinas is not consistent in his writings in what role he ascribes to particular internal senses. In some cases he follows the terminology of Aristotle, in other cases the two philosophers differ to a considerable degree. We can observe a great confusion with respect, for instance, to Aristotle's and Aquinas's treatment of the internal sense of *phantasia*. As we shall see, it is difficult to figure out in the first place what role Aristotle wanted to ascribe to this internal sense. Some passages suggest that *phantasia* for Aristotle is a passive faculty to experience appearances; others, that it is more like an active faculty of imagination (see below 2.3.3). In any case, *phantasia* seems to be playing a crucial role in Aristotle's account of cognition. Aristotle says at some point that "there is no supposition without it" (427b15) and treats *phantasia* as a necessary condition of thought. What is even more interesting, however, is that it appears that Aristotle's *phantasia* could plausibly be interpreted as a quite sophisticated *innate cognitive mechanism*. Aristotle, like Aquinas, is typically considered a cognitive empiricist. This conviction, however, should be revised once we carefully analyze Aristotle's account of the internal sense of *phantasia*.

As concerns Aquinas, in the *Summa Theologiae* the internal sense of *phantasia* is identical to imagination. The function that Aquinas ascribes to *phantasia*-imagination in the *Summa Theologiae* is, however, surprisingly limited: it plays the role of memory, i.e. it is a treasury for sensible species. In addition, in his *InDA* Aquinas does not treat *phantasia* as a separate faculty but rather uses it "as a generic term covering three distinct faculties of internal sense: the imagination, the *vis cogitativa* and the sense memory"
In fact, it is the *vis cogitativa* (the cogitative power) that is the most important internal sense in Aquinas's account of cognition.

**2.3.2 The internal vs. the external senses**

*(i) Passivity and Materiality*

According to Aquinas, a characteristic shared by all of the senses is that their functioning is dependent on specific bodily organs. The external and the internal senses differ, however, in that whereas each of the external senses has an 'external' part of the body as its organ, all the internal senses have their organs in the brain (*DV* 18, 8). Furthermore, for Aquinas all of the senses, as opposed to the intellect, are *material*. Their materiality, however, should not be confused with the requirement that each of the senses should be assigned to a specific bodily organ. As we said above, the terms 'materiality' and 'immateriality' in Aquinas's account of cognition refer rather to the degree of potentiality of any given cognitive faculty (cf. above, 1.1.1). We'd say, therefore, that, on Aquinas's account, the internal senses are *less material* than the external senses because the set of objects that they can cognize is less limited (it is to a lesser degree determined by the objects' physical features).

Finally, Aquinas characterizes the external senses as "passive capacities" that are "naturally suited to be impressed upon by an external sensible quality" (*ST* I 78, 3). The external senses are passive in that they take no active part in creation of their proper objects. They don't need to be active because, for Aquinas, the material objects in the
external world are themselves actually sensible and ready for being sensed: the material objects impress their forms on the bodily organs of the external sense organs. All the information that the external senses process (all the content that will be represented by the sensible species) comes from the external world. No additional information comes from the external senses. The role of these senses is just to receive the forms from the external objects, and to translate the information received (physical stimuli) into language that will be understandable for cognitive faculties of higher levels (and so, into neural signals).

The situation is somewhat different with respect to the internal senses. They demonstrate a similarity with the intellect in that they to some extent also create their proper object, that is, phantasms. As we said, although phantasms are produced on the basis of the data processed by the external senses and contained in the sensible species, they also transcend this data; phantasms contain something more than sensible species. The internal senses, nonetheless, still count as passive because they cannot put themselves into action. Their activity is automatically initiated by means of efficient causation by the working of the external senses. (As concerns the intellect, Aquinas would say that it is entirely active in its character: because it is immaterial, it cannot in any way be affected by phantasms. On the contrary, it is the agent intellect that takes the active role and efficiently acts on phantasms. It seems, nonetheless, that phantasms must

---

71 This distinguishes the senses from the intellect: because things in the external world are not actually intelligible, there must be an active element in the intellect.
have some influence on the intellect—perhaps they trigger the intellectual activity somehow—otherwise, objectivity of cognition would be threatened.)

(iii) Different Formal Objects

The clearest difference between the external and the internal senses concerns their respective proper objects (see 1.2.3 (iv) and (v)). The external senses directly cognize (and the sensible species represent) proper (color, sound, smell, etc.) and common sensibles (size, number, shape, etc.). The proper objects of the internal senses are incidental sensibles, or intentiones non-sensatae, i.e., forms which are not directly perceived by the external senses (ST I 78, 4). Aquinas explains that

Q45. for something to be sensible per accidens, the first thing that is required is that it be an accident of something sensible per se. For example, being a human being applies accidentally to what is white.

A form can be a proper object of the internal senses only in virtue of being an accidental form of something that is sensible per se. So, for instance, 'black' is a color, and color is the proper object of the sense of sight: it is, therefore, sensible per se. A cat can be a proper object of the internal senses in virtue of being an individual substance which 'applies accidentally to what is black.' And so, the proper objects of the internal senses turn out to be the same thing as the incidental objects of the external senses, or, in other words, as sensible per accidens. This shows a close dependence of phantasms on sensible species.
(iii) The Intelligible Contained in the Sensible

It is worth noting that the incidental sensibles or intentiones non-sensatae, which, according to Aquinas, are cognized per se by the estimative and cogitative powers, are also sometimes understood as that which is "intelligible in that which is sensed" (Klubertanz, 1952, p. 196). This suggests that some intelligible element is already somehow contained in the sensible species and that it is accessible to the internal senses.

What constitutes a great challenge is the sense in which the intelligible is supposed to be contained in the sensible. The FT holds that it is literally there; it is enough just to ignore the sensible in order to be able to abstract the intelligible. The FTF, on the other hand, understands this phrase differently: the intelligible is already there in the sensible only in the sense that there is a potential in the cognitive powers to read it off. There is some kind of innate mechanism at the higher level of cognition that knows what to do with the data provided by the external senses. The following passage confirms this interpretation:

Q46. Per accidens that is sensed which does not affect the sense inasmuch as it is a sense, nor as it is this sense, but as joined to those things which of themselves affect the sense, as 'Socrates,' and 'the son of Diaries,' and 'friend,' and other like things. These things are known in the universal by the intellect; in the particular, they are known by the discursive power in men, and by the estimative in other animals. Such things the exterior sense is said to sense, even though only per accidens. (In IV Sent 49, 2)

What this shows is, not surprisingly, that the external senses do not have a direct access to what is intelligible. At the same time, in every normal, non-defective occurrence of the human cognition, there is a connection, a non-direct link between the external senses and the intelligible. The proper functioning of the external senses in human beings
immediately indicates, even though only *per accidens*, the presence of something intelligible. And so, the above passage from *In IV Sent* 49, 2 also shows that for Aquinas that which is intelligible in things is accessible not only to the intellect with its universal cognition, but also to some cognitive faculty at the level of the internal senses. As we shall see, it is the role of the cogitative power to take the data provided by the external senses and to put it together in the right way, to transform it in such a way that the subject can determinately and accurately apprehend a particular concrete object and to cognize it as an object of a kind.

(iv) Lower-Level Sensation

According to Aquinas both rational and non-rational animals have a set of natural inclinations, or natural appetites, and a set of sensory appetites, both determined by the animal's substantial form. The function of an appetitive power is to produce desires which in turn lead to action (*ST I-II* 1, 2, *SCG* III 2-3); as Aquinas puts it, "every animal that moves does so in order to pursue something desired or intended" (*ST I* 78, 1).

Natural appetites are the most primitive and predictable. They are the only inclinations possessed by such beings as plants. (Think, for instance, of your indoor plants' tendency to lean over towards the window so that they face the light.) In animals, natural appetites are at work in situations in which an animal reacts to stimuli in an automatic (reflexive) way, independently of any (higher kind of) cognition, and where (in principle) no modification of behavior and no learning is possible. Then, the sensitive soul endows organisms also with sensory appetite. As Pasnau explains it, "animals, in
virtue of this appetite, have the capacity to modify their desires in light of changing circumstances" (Pasnau, 2002, p. 209). What this means is that the beings endowed with the sensitive soul are not as determined in their behavior as were the vegetative and lower existing things. Animals can learn to react differently to the same stimulus in different circumstances. If a caribou encounters an electric fence on the path it usually follows to the stream, it will look for another way to get to the source of water: it does not have to stick to the same path. (It is not so with plants: plants will continue to grow towards light no matter what, even if the light is going to burn them.) Animals, moreover, do not just passively react to stimuli coming from the external world. They can learn that changes in their own behavior may bring about desirable events. A monkey, for instance, can figure out that by intimidating another monkey, it will be able to take its food.

And so, we can see that while desires lead to action, they are not produced at random, but rather in response to some information that has been received from the environment (ST I 78, 4 and 80, 1, ad3). Thus, Aquinas's account can be summarized as follows: every action follows appetite, and all appetite is consequent upon apprehension or cognition. This already indicates that the crucial factor for the animal's chances for survival is its cognitive endowment (cf. 5.2.5.)

(v) Instinct

The apprehension that influences the appetites of non-rational animals is twofold. One part of it is based on the information obtained by the external senses, the other, on the
data provided by the internal senses. The behavior in which an animal seeks what is pleasant and avoids what is painful is determined by the external senses (it is caused by the stimuli received by the sense organs). The external senses, however, are limited to what is actually present, or to what is presently sensed. Animals, on the other hand, must also perform actions "with respect to objects that apparently cause no immediate sensory pleasure or displeasure" (Klubertanz, 1952, p. 227). Yellow-bellied marmots, for instance, search for seeds in preparation for winter hibernation already in the late summer. In order for these kinds of behavior to be possible animals need to "have a power of apprehension which goes beyond the special senses, and consequently beyond the retention of the knowledge thus acquired" (Klubertanz, 1952, p. 230): this is what is needed for the animal to survive.

Non-rational animals need to be able to distinguish what is dangerous for them from what is beneficial, what should be pursued from what should be avoided. Aquinas explains:

Q47. If an animal were moved by pleasing and disagreeable things only as affecting the sense, there would be no need to suppose that an animal has a power besides the apprehension of those forms which the senses perceive, and in which the animal takes pleasure, or from which it shrinks with horror. But the animal needs to seek or to avoid certain things, not only because they are pleasing or otherwise to the senses, but also on account of other advantages and uses, or disadvantages: just as the sheep runs away when it sees a wolf, not on account of its color or shape, but [because it perceives it] as a natural enemy [*sed quasi inimicus naturae*]; and again a bird gathers together straws, not because they are pleasant to the sense, but because they are useful for building its nest. Animals, therefore, need to perceive such intentions, which the exterior sense does not perceive. And some distinct principle is necessary for this; since the perception of sensible forms comes by an immutation caused by the sensible, which is not the case with the perception of those intentions. (*ST* I 78, 4)
This function of going beyond and completing the information provided by the external senses is performed in non-rational animals by their natural instinct, or, as Aquinas calls it, by the internal sense faculty of the estimative power.²² (In human beings, the vis cogitativa, or the cogitative power, is a counterpart of this natural instinct.)

Instinct, Aquinas would say, is a natural judging power indicating the proper good for a given being. It is necessary for survival and for normal functioning of an animal. Instinct ‘reads off’ reality what the external senses by themselves cannot see. For instance, the color and shape of a wolf are not enough to see in it an enemy. Rather, the capacity to recognize wolves as something that needs to be fled must be somehow ‘built into’ the estimative power of animals of some species.²³ The estimative power can be explained, therefore, as an innate mechanism thanks to which a certain specific set of sensible species gives rise, for instance, in sheep, to a phantasm containing information ‘enemy.’ Another set of sensible species would give rise, in sheep, to a phantasm ‘food,’ and yet another — to a phantasm ‘offspring to be fed.’

Q48. Thus a sheep knows this particular lamb, not as this lamb, but simply as something to be suckled; and it knows this grass just in so far as this grass is its food. Hence, other individual things which have no relation to its own actions or passions it does not apprehend at all by natural instinct. For the purpose of natural instinct in animals is to direct them in their actions and passions, so as to seek and avoid things according to the requirements of their nature. (InDA II 13, 398)

For Aquinas, the function of instinct in general is threefold. It is responsible for:

a) the preservation of life,

---

²² Aquinas also talks about the human estimative power. This cognitive faculty only works in exceptional situations: when there is something wrong with the intellect. Cf. InDA III 6, 669-70.

b) reproduction, and

c) cognition.\textsuperscript{74}

Because it is directed at the well-being of the organism, instinct can be called, first of all, an innate tendency to preserve life. The natural impulse to live does not seem to be optional.\textsuperscript{75} Every animate thing (with rare exceptions, such as, e.g., a person who decides to become a kamikaze) wants to live and strives to prolong its own life. Connected to it are the natural impulse to reproduce, to preserve one's species, and the tendency to cognize. Reproduction serves preservation of life. Cognition serves both preservation of life and reproduction. Here again it is not optional: an animal cannot help it, it \textit{has to} register information about the environment. The sheep will not run away from the wolf unless it perceives it; but also, it will not run away from it unless it perceives it as an enemy. Similarly, the species will not survive unless animals can recognize their mating partners.

Obviously, human and animal instincts differ in various respects.\textsuperscript{76} For instance, in what concerns the instinct to preserve life, we lack the instinctive knowledge that animals have of what they should or should not eat, or what is good for them, and what

\textsuperscript{74} Even though it makes sense to talk of the animal instinct as having these three functions, the first function, the preservation of life really seems crucial. This is the function that is served in some sense by all faculties (both cognitive and non-cognitive) in an animal.

\textsuperscript{75} Aquinas would use the word 'desire' in this context: he'd say that all animate beings have a \textit{natural desire} to preserve their lives. In our times, however, the term 'desire' brings to mind something optional. So, when we hear Aquinas talking about a 'natural desire to preserve one's life,' we should translate it into innate instinct, reflex, or impulse (where it does \textit{not} have the sense of a mental state). We could talk, for instance, of the natural impulse to avoid pain, or to seek pleasure. A natural reaction is to pull your hand (or your paw) away from fire when you start feeling pain—in this sense it is not optional; even though humans can and sometimes do decide to oppose or distort their natural impulses of this kind.

they should avoid. In addition, while following natural instinct is not optional for animals, in humans it is or can be under the control of reason. Both humans and non-rational animals have an innate instinctive tendency to copulate in order to satisfy their sexual desire, and to eat in order to satisfy the feeling of hunger. In humans, however, are able to go against such instincts—they can fast or remain celibate.

Still, what is crucial for our considerations is that in both non-rational animals and in humans there are these instincts: instinctive impulses and tendencies, and also instinctive mechanisms that facilitate the fulfillment of these tendencies. Cognition (including the lower level of intellectual cognition) is also one of such instincts. It is a capacity with respect to which humans and non-rational animals differ in important ways. In both cases, however, cognition is something instinctive; it is a manifestation of one aspect of the natural endowment that every living creature possesses. That animal cognition is something instinctive sounds pretty commonsensical. It is somewhat more controversial to claim that human cognition is so as well. It should make more sense, however, if we emphasize that something, e.g., a mechanism or a behavior, is

---

77 In humans, as it seems, these two innate tendencies have been severed, to a greater or a lesser extent, from what seemed to be their natural purpose, namely, reproduction and the sustaining of life. That sexual desire is severed from its reproductive function can be seen in the popularity of contraception; the couch-potato syndrome suggests that the desire for food is similarly severed from its function to sustain one's life. Maybe both have been replaced by a super-instinct to feel or pursue pleasure. Unfortunately, this one does not necessarily serve the preservation of the species—as, for instance, many European nations are currently able to verify. With respect to the innate tendency to reproduce, this in turn seems to have been replaced by what we may call a new 'instinct' (or incontrollable impulse) to make money. (The way in which people live in our times suggests that indeed they are genetically programmed not to reproduce, but to make money.)

78 It is sometimes said that even the most basic instincts in humans can be controlled and even gotten rid of. If this was the case, then obviously it would not make sense to talk about an 'instinct to use concepts' or an 'instinct to cognize the universe as containing individual beings.' We certainly cannot decide one day not to use concepts any more. But then, perhaps the innate character of cognition goes even deeper than instinct; perhaps it is more like digestion in that it is really not in our control.
'instinctive' in the sense of being 'innate.' Cognition is innate and instinct-like, as we are going to show, in how it works, and also because of how it developed. It is instinct-like in that it consists of a set of innate mechanisms that enhance the chances of survival for a given species. And it is instinct-like because it is a product of natural selection. Its functioning, as the functioning of all instinctive or innate behaviors, is an evolutionary adaptation for the human species.

The most important goal of the human instinct for cognition is the same as the purpose of instinct in general in all animals: for a given organism to have a better chance to survive. Instinct enables non-rational animals to cognize what is necessary and useful for their individual and specific nature, or what should be avoided because it is dangerous or detrimental. This kind of instinct is missing in humans. Humans are not only limited in their sensory cognition (their external senses are less capable), but also natural instinct does not immediately provide them with the knowledge of what is necessary for survival and what is detrimental (even if, to some extent, we might share with cats the innate fear of heights, still we smoke cigarettes, and drink neutral grain spirits—things that no instinct-following non-rational animal would ever do). But, according to Aquinas, "nature does not fail in necessary things" (ST I 78, 4). And so, while humans do not have the same kind of survival-enhancing natural instinct as do non-rational animals, the human cognitive instinct consists instead, as we shall see
below, in an instinctive tendency\textsuperscript{79} to cognize things as individual substances, and as things of a certain \textit{kind}, falling under a universal concept. This instinct allows for a better understanding and control over the world, and increases our chances of survival.

Before we discuss Aquinas's views on the cogitative power, we will have a look at Aristotle's view of the internal sense of \textit{phantasia} that also can be interpreted as a quite sophisticated innate element, an innate cognitive mechanism.

### 2.3.3 Aristotle on \textit{phantasia}\textsuperscript{80}

(i) \textit{Phantasia} in Aristotle's \textit{De Anima}

Aristotle presents his views on \textit{phantasia} in book III of \textit{De Anima}. \textit{Phantasia}, for Aristotle, is "a movement resulting from an actual exercise of a power of sense" (429a1-2), it is "impossible without sensation," it occurs in beings that are percipient" and it "has for its content what can be perceived" (428b11-3). Phantasms, on Aristotle's account, are mental images that "remain in the organs of sense and resemble sensation" (429a6). They can be understood as various kinds of retained sensory impressions: "each of them is (...) the remnant of a sensory impression taken when sense was activating itself" (\textit{On Dreams} 461b22-25). As Aquinas clarifies, phantasms "'dwell within' in the absence of sensible objects, as traces of actual sensations." (\textit{InDA} III 3, 669).

\textsuperscript{79} When we use the term 'instinct' with reference to cognition, it has a somewhat different meaning from the more common use of the term. It is not what Aristotle has in mind when he says, e.g., that the role of human reason is to bring our 'instinctive' behavior under control; cf. \textit{Nicomachean Ethics}, e.g., Book I.

\textsuperscript{80} When I analyze Aristotle's texts, I will often refer to Aquinas's suggestions on how they should be interpreted. The outcome, therefore, will be something like 'Aristotle read through the eyes of Aquinas.'
Phantasia in Aristotle's texts is usually translated as either (1) appearance, or (2) imagination. This suggests two possible interpretations: on the one hand, phantasia can be seen as a passive faculty to experience appearances, a faculty that allows for something to happen to the subject, as when an object appears in front of his eyes. On the other hand, phantasia as imagination would be an active faculty which allows the subject to control whether some thing will appear to him in his mind, as when I make myself imagine what Pegasus would look like. In both cases phantasia is taken to cover one of the following three things: "it designates the capacity, the activity or process, and the product, or result" (Frede, 1992, p. 279). Thus, phantasia as appearance can mean either (i) the capacity to experience an appearance, (ii) the on-going appearance itself, or (iii) what appears. Similarly, phantasia as imagination means (i) the capacity to create appearances, or images, (ii) the on-going creation of images, or (iii) created appearances. I am going to focus on option (i) in both cases.

(ii) Phantasia as Imagination

Based on Aristotle's texts, we can define the faculty of phantasia understood as imagination by means of the following features: (1) it is an internal sense faculty operating on phantasms; (2) it is at work (usually) when the external senses are not, that is, when the sensible object is not present; (3) except when it is operating during sleep, it can be controlled, that is, it can be activated at will by the cognizing subject, and finally, (4) it (usually) offers an imperfect or even distorted picture of the cognized object.
Phantasia as imagination, similarly as other internal senses, is a faculty operating on phantasms. And like other internal senses it is also strictly dependent on the external senses: phantasms are produced on the basis of sensible species, and so, without prior activation of the external senses there would be no phantasms at all. On the understanding of phantasia as imagination, its role is to choose among sensible forms stored in memory, and, on this basis, to produce images of sensible objects. To give an example, we can say that when I went jogging in the morning, I acquired a sensible form of my neighbor's dog by means of a direct sensory experience: I saw something black, small, with floppy ears, etc. Now the dog is no longer available to my senses: she is nowhere to be seen or heard. However, I can still recall the image of her in front of my eyes: I can imagine her chasing a cat, or how she would look if my niece put a red dress on her. Similarly, it is the internal sense of phantasia as imagination that is at work when I recall last summer's vacation in the mountains, or try to imagine what will happen in the future.

Phantasia as imagination can be activated when the sensible objects are not present in front of the perceiver, that is, when the objects of some past occurrence of sense-perception are already gone but their forms, or sensible species, are remembered, or in a situation in which the objects imagined were never even present to the perceiver, as it happens, for instance, when I imagine a centaur. The idea here is that I can be properly said to imagine a horse only when the external senses are not currently being

---

81 This may suggest that phantasia is similar to memory. The relevant difference is that phantasia as imagination is an active faculty: it allows us to create images of things that we have never seen before.
affected by a real horse in the external world. A similar situation happens, according to Aristotle, also when we dream, and so, on Aristotle’s account, it is the faculty of phantasia that allows us to see or hear things while we sleep.

According to Aristotle,

Q49. a man can exercise his knowledge when he wishes, but his sensation does not depend upon himself—a sensible species must be there. (DA 417b23-25)

We have no control over our sensation: unless I hallucinate, I can see a black dog in front of me when she really is there, and only if she really is there. Once the dog is gone, I do not see her any more. The image of the dog, however, might still be in my memory: I can easily imagine her, that is, I can create or recall an image of how she looks. In fact, I can imagine almost anything I wish: a yellow dog with horse’s ears, for instance.

Thus, phantasia as imagination, with the exception of dreams, is up to us. As Aristotle puts it, "imagining lies within our own power whenever we wish (e.g. we can call up a picture, as in the practice of mnemonics by the use of mental images)" (DA 427b18-21). Interestingly, this feature makes the working of phantasia less similar to that of the external senses, and closer to that of the intellect: both phantasia as imagination and the intellect can be controlled by the cognizing subject.

Another characteristic of phantasia as imagination, this time more surprising, concerns its lack of reliability with respect to representing the external world. The function of every faculty that plays a role in cognition seems to be to provide the subject...
with a correct representation of the objects in the external world. However, this is not the case with respect to *phantasia* as imagination. Aristotle says that while "sensations are always true, imaginations are for the most part false" (*DA* 428b11-12). Here, therefore, the products of *phantasia* appear to be something like unclear or distorted perceptions: they are 'false' in the sense that they do not correspond to anything in the extra-mental world. This can be seen, e.g., in the following passage from *On Dreams*:

Q50. just as in a liquid, if one vehemently disturbs it, sometimes no reflected image appears, while at other times one appears, indeed, but utterly distorted, so as to seem quite unlike its original; while, when once the motion has ceased, the reflected images are clear and plain; in the same manner during sleep the phantasms, or residuary movements, which are based upon the sensory impressions, become sometimes quite obliterated by the above described motion when too violent; while at other times the sights are indeed seen, but confused and weird. (*On Dreams* 461a15-23)

(iii) Problems with the Interpretation of *Phantasia* as Imagination

Given the importance Aristotle ascribes to *phantasia* in both sensory and intellectual cognition (he says that "there is no supposition without it" (427b15) and treats *phantasia* as a necessary condition of thought; cf. Frede, 1992, p. 280), it is not clear that the above interpretation would be appropriate. The understanding of *phantasia* as imagination does not provide us with a good account of the difference between phantasms and the corresponding sense-perceptions. Nor does it explain in what sense *phantasia* is a mediating stage in the process of cognition between the external senses and the intellect. Most importantly, *phantasia* as imagination does not seem to be a good candidate for what is supposed to be a necessary condition for thought: it is not plausible to claim that
something that is usually confused, plainly false, or that provides a distorted image of reality, could be at the same time what the intellect requires for cognition.

So perhaps what we deal with here are two distinct faculties. What we describe as imagination, that is, this active faculty that can be controlled by the subject, is not the same thing as the faculty of phantasia that is required for the intellectual cognition. In what follows, therefore, we are going to treat imagination as a separate faculty, not identical to phantasia as appearance. The latter, from now on, I am going to call phantasia tout court.

(iv) Phantasia as Appearance

On Aristotle's account, phantasia is the highest-level cognitive faculty for all non-rational animals. As it is the case with Aquinas, also for Aristotle, the external senses together with the common sense are not sufficient for cognition in the genuine meaning of the term. In Aristotle's case, it is only phantasia as appearance that makes cognition possible.

There are various places in Aristotle's texts which suggest, contrary to what has been said above, that phantasia is at work also during actual sensations. In book III of De Anima Aristotle says, for instance, that the acts of phantasia concerning proper objects of sense are "free from error while the sensation is present," and that those acts of phantasia that concern common or incidental objects of sense "may be erroneous whether it is present or absent, especially when the object of perception is far off" (DA 428b27-30). Moreover, Aristotle explains that phantasia is "impossible without sensation," and "is
necessarily similar in character to the sensation itself" (DA 428b12-14). This similarity can be explained by the fact that for every act of phantasia, there is a corresponding act of sensation, and vice versa. The first part of this claim should not be surprising: the same holds also with respect to the faculty of imagination. For any act of imagination, there must be a corresponding act of sensation (or several acts) that has happened in the past: I can only imagine a beautiful horse because (in typical cases) I have seen different horses in the past. Now, with what concerns phantasia as the faculty of appearance, we can make an even stronger claim, namely, that the relationship holds both ways: in any non-defective cognizing subject, for any act of sensation, if it is to be able to play a role in cognition, that is, if it is to affect what the subject cognizes, there must be a corresponding act of phantasia.

So, the situation, right now, is as follows: while the reception of sensible species is possible only when the object of cognition is present, the acts of phantasia, and the processing of phantasms in general, may take place both when the object of cognition is present and when it is not. This may seem redundant: the operation of phantasia turns out to be, at least at times, simultaneous with the working of the external senses. It is not clear, therefore, why phantasia is supposed to be necessary for cognition even when the external senses are at work.

On both Aristotle's and Aquinas's account of cognition, the external senses together with the common sense allow us to cognize things only while perception actually occurs, only when the sensible object is in front of us. I hear thunder when the sound of it causes the vibrations in my inner ear (when the vibration is actually
stimulating my organ of hearing). Now, consider what must happen for us to be able to understand a whole sentence. On what we might call a purely physiological level, we register each particular word in succession, just one sound at any given moment. That is to say, the stimulus that affects our sense of hearing is something discrete. In order to understand a sentence, however, it is required that the words of the sentence somehow form a continuum. On Aristotle's account, it is the internal sense of phantasia that enables us to 'still hear' the beginning of the sentence in our minds when what we actually hear, that is, the sound that is actually causing vibrations of our inner ear, is the end of it, and thus to understand a long string of words as a sentence, and not as a discrete succession of sounds. Similarly, the ear—an external sense organ—registers different vibrations; but it is phantasia that allows us to hear a melody.

Dorothea Frede illustrates the same idea with the example of the sense of sight:

Q51. Because of the emphasis on the singleness of each act of perception and on the need for the presence of its object, it is doubtful that for Aristotle we can have something like a 'panoramic' view of a whole situation, for he does not seem to include anything like a 'field of vision' in his explanations. This would suggest that when I let my eyes glide over the different books on my bookshelves there is always just the piecemeal vision of this or that colored object; the overall impression of the different books (including those behind my back) would then be already a phantasia, a synthesis of what I perceive right now and what I perceived a second ago and so on. (Frede, 1992, p. 283)

On this interpretation, therefore, phantasia can be said to transcend physiological limitations of the external senses. Phantasia is able to play this role because, as we have said, phantasms can be operative also when the sensible object is gone, that is, also when the stimulus is not actually affecting our external senses. As Frede explains,
Q52. [phantasms] can be separated from their origin, while perception cannot, and this means that they can give us a coherent picture of a situation that transcends the immediate perception. [Phantasia] can give us the impression of a change over a certain time, as when my eyes glide over different objects in this room or my ears follow a melody. Strictly speaking, the eyes or ears perceive only one object at a time; thus animals without phantasia would only get a sequence of incoherent imprints. (Frede, 1992, p. 285)

One could argue that the internal sense of memory could be sufficient to play the role that we want to ascribe to phantasia. Most likely memory, as a faculty which retains and preserves the sensible species, is indeed involved in the working of phantasia. Memory by itself, however, would not be sufficient to play the role that we ascribe to phantasia: we need a faculty that is able to apprehend at the same time both those sensible species that already belong to the past (the beginning of the sentence that we heard a moment ago), and those that are affecting our external senses at the present moment.

The common sense, similarly, could not play the role that we ascribe to phantasia because the common sense "qua sense contains no more than the imprints of the different sense-perceptions at any moment" (Frede, 1992, p. 284). Thus, Frede concludes, "since even in the [common] sense the imprints of the perceptible forms last only as long as the perception itself, what lingers on in it when I avert my eye is then already a phantasia, an after-image" (Frede, 1992, p. 284). Sensation, that is, the external senses together with the common sense but without any involvement of phantasia, is limited to what is perceived at the moment of simultaneous contact of the senses with the stimuli coming from the outside world, exactly when they affect the organs of the external senses.

Notice that on this interpretation, the faculty of phantasia would not be exclusive to humans. On the contrary, it would be necessary to ascribe this faculty also to many
non-rational animals. Consider, for instance, birds recognizing a bird-call of many notes, i.e., a musical phrase. Similarly as it enables us to hear a whole sentence instead of a sequence of separate word-sounds, *phantasia* allows a bird to hear a sequence of sounds as one continuous melody.

This, I believe, is what makes the whole thing even more plausible. Without an account of the kind of internal sense similar to *phantasia*, it would be difficult, if not plainly impossible, to ascribe any kind of cognition to animals that do not possess the intellect. They would be limited in their contact with the external world to the information provided by the external senses, perhaps together with the common sense. The external senses, however, are not enough for cognition. It is the internal sense of *phantasia* that, on Aristotle's account, provides an explanation of cognition of a lower kind in which there is no involvement of the intellect, but for which the internal senses are necessary in addition to the external senses.

On Aristotle's account, the role of the internal sense of *phantasia* is to go beyond the external senses and to cognize something that was not yet in the sensible species. *Phantasia*’s functioning suggests that it could be interpreted as instinctive, innate mechanism, useful for the species. Such a mechanism, we would say today, is a product of natural selection; it is fixed in a given species endowed with cognition because of its survival value. This shows, against most common interpretations, that Aristotle is not a pure empiricist about cognition. The nativism that we find in his treatment of *phantasia* is more interesting and more sophisticated than the most primitive kind of nativism accepted even by the hardest among the empiricists: the kind of nativism which posits
the most simple innate mechanisms that register stimuli incoming from the external world and translate them into language that can be understood by the higher cognitive faculties.

2.3.4 Aquinas on the *vis cogitativa*

On Aquinas's account, the counterpart of the animal instinct, that is, of the animal estimative power, is the cogitative power—an internal sense unique to human beings. The two powers are to some extent similar: for instance, the proper objects of both are *intentiones non-sensatae*. The estimative power triggers instinctive reactions in non-rational animals, the cogitative power also governs human instinctive behaviors. The cogitative power is *the sole* governor of the human behavior, however, only "when the intellect is not in charge (...) as the result of some passion (one of anger, concupiscence, fear, or any such thing) that is so strong as to cloud intellect," "as the result of some illness" (*InDA* III 6, 669-70), or else in newborn babies. In these situations, it is better to refer to this power as the 'human estimative.' In all other cases of 'normal' cognition, the cogitative power in human beings always works in collaboration with the intellect.

In non-rational animals the natural instinct, i.e., the *vis estimativa*, is sufficiently developed to guarantee that the animal possesses those capacities that are necessary for the preservation of the individual or the whole species. The human instinct, on the other

---

83 In the passage from *InDA* III 6, 669–70 Aquinas uses the term 'phantasia' to refer to the faculty governing instinctive behavior. What he means, however, is more likely estimative and cognitive powers. *Phantasia* for Aquinas is not a cognitive power, and for this reason it could not govern actions.
hand, is in a way incomplete, or underdeveloped (ST I-II 77, 1). It needs the aid of the intellect to be able to discern what is truly good for the human being. The intellect is supposed to control and often go against our instinctive impulses. For this reason some of Aquinas scholars suggest that animal and human instincts are entirely different:

Q53. In animals this instinct, or estimative power, is absolutely perfect and unerring. (...) Our [human] instinct (...) is much less perfect than that of animals, and so we are forced to rely on our intellect. (Swiezawski, 1995, p. 137)

This interpretation, however, does not provide a complete picture of Aquinas's views. Rather, it seems more useful to distinguish two meanings of the term 'instinct' when talking about human beings. On the one hand, instinct can be considered as a more limited faculty, when it refers to some inborn action patterns in the domains of sexuality, reproduction, self-preservation, aggression, etc. In this sense it can indeed be said that the role of the human intellect is to control instinctive impulses. On the other hand, however, a plausible way to interpret Aquinas's views is that while human beings lack the kind of natural instinct that would unmistakably recognize "the usefulness or harmfulness of certain perceptions" (Swiezawski, 1995, p. 137), they do possess another, and in fact, a more powerful kind of instinct which is the cogitative power.

The vis cogitativa appears in Aquinas's writings under two additional names: it is called both (1) particular reason and (2) the human instinct. Supposedly, these two

---

84 Swiezawski continues (p. 137): "Animals, by virtue of their estimative power, act unerringly. They know hic et nunc—in any given time, place, or circumstance—what is absolutely necessary for them. They select exactly the objects they need. We see this absolute purposefulness and flawlessness everywhere—in the migration of birds and fish, in the building of nests by birds and the construction of dams by beavers. The same applies to the things animals avoid. Instinct gives them the invaluable ability to discriminate between what is beneficial and harmful."
names indicate merely that the cogitative power is at the same time sensory (and so it is close to the highest cognitive power of those beings whose cognitive capacities do not go beyond the sensory, i.e., to the estimative power of non-rational animals) and intellective (it is called 'particular reason' because it is capable of some sort of categorization). My claim is that in addition the two names of the *vis cogitativa* indicate a specific kind of innateness of this power.

As various passages in Aquinas's writings suggest, the two main functions of the cogitative power are to enable a human cognizing subject

(a) to be aware of an individual as an individual and
(b) to recognize an individual as a member of a certain kind.

These two functions, which place the cogitative power in the intermediary position between the intellect and the external senses, allow it to play a crucial role in explaining the relationship between sensation and cognition.

**(i) Cognition of Individuals as Individuals**

In his commentary on Aristotle's *De Anima* Aquinas characterizes the internal sense of the *vis cogitativa* as follows:

Q54. [when an object] is apprehended as an individual—e.g., when I see something colored I perceive this human being or this animal—then this sort of apprehension in a human being is produced through the cogitative power. This is also called particular reason, because it joins individual intentions in the way that universal reason joins universal concepts. But all the same, this power is in the soul's sensory part. For the sensory power, at its highest level, participates somewhat in the intellective power in a human being, in whom sense is connected to intellect. (*InDA* II 13, 397–8)
The first function of the *vis cogitativa* is to apprehend individual objects. This, in fact, is a crucial element in Aquinas's account of cognition, because it allows the cogitative power to fill the gap between the external senses and the intellect. As we have said above, the external senses together with the common sense provide us with knowledge of accidental features of an object. On the other hand, the primary function of the intellect is to cognize essences and to provide the subject with universal knowledge. The intellect, therefore, does not (or, at least, not directly) cognize *things*. Essences that it does cognize must be separated from all material conditions; they are not individual but common to many things (cf. *ST* I 3, 3). What this shows is that the external senses together with the intellect still do not allow the cognizing subject to reach the principal ontological category in Aquinas's doctrine, namely, primary substances, that is, things like human persons, or cats, or stones.

Imagine there is a cat in front of John. The external senses tell John that there is in front of him something gray, furry, stretching and meowing. Assuming that John already has the concept CAT, his intellect would inform him about the content of the concept, about what constitutes the essence of cats. But it is only thanks to the cogitative power that John, on the basis of the data registered by the external senses, is able to focus on the cat (and not, e.g., on its tail together with a piece of carpet on which it stands), and to cognize it as *this individual cat*. 

140
(ii) Cognition of Individuals as Members of Various Kinds

Aquinas also says that "the cogitative power apprehends an individual as existing under a common nature. (...) it cognizes this human being as it is this human being, and this piece of wood as it is this piece of wood" (InDA II 13, 398). This suggests that the cogitative power must possess or at least somehow be able to apply the appropriate categories of 'human being' and 'wood.' The cogitative power, however, because it is a sensory power, can be said to possess or apply categories only because it collaborates with the intellect.

Aquinas explains that the *vis cogitativa* "is on the boundary of the sensitive and intellective parts, where the sensitive part touches the intellective" (In III Sent 23, 2, 2). The difference between the intellect and the cogitative power consists in their proper objects: whereas the intellect apprehends universal concepts, the cogitative power operates on individual intentions, i.e. on forms which represent only individual objects.\(^85\) This way Aquinas makes sure that the *vis cogitativa* is still a sensory power: an intellective power on its own could not operate on forms of this kind.

At the same time, Aquinas emphasizes that the *vis cogitativa* significantly differs from the external senses. He says that the cogitative power is "united to the intellective power in the same subject" (InDA II 13, 398), and that it is *superior* to the external senses "not through something proper to the sensory part, but through a kind of affinity and

---

\(^85\) It seems that on Aquinas's account all concepts are universal. For a concept to be individual, it must be somehow sensory. On such an account, the concept GOD, for instance, does not count as individual. Properly speaking, the concept GOD is a concept of a genus, which just happens to be instantiated by one object.
closeness to universal reason” (ST I 78, 4, ad5). The intellect and the cogitative power are indeed similar in various respects. One concerns the way in which they operate. The *vis cogitativa*, that Aquinas also calls 'the particular reason' (ST I 78, 4), "joins and divides individual intentional likenesses" (DV 10, 5). This corresponds to the second operation of the intellect, *compositio* and *divisio* in which it "puts together or takes apart" universal concepts (cf. above, 1.1.2 (v)). In fact, in ST I 78, 4 Aquinas lists "adding, composing, and dividing" as characteristic acts of the cogitative power that distinguish it from the estimative power. By operating on phantasms, the cogitative power has access to individual things in the external world. Because of its cooperation with the intellect, on the other hand, it can be aware of an individual as a member of a particular kind, that is, it can assign sensed primary substances to broader classes. The cogitative power can see things as belonging to different categories; it can also make comparisons between individual objects (cf. Pasnau, 2002, pp. 253–5), and, plausibly, even make inferences involving these objects.

It is far from clear in what way the intellect is supposed to supply the cogitative power with basic conceptual categories and whether, as a result, the cogitative power should be said to possess universal concepts. Aquinas himself does not provide any details about how the collaboration between the intellect and the cogitative power is supposed to work. What we can figure out based on his account, however, is the character of the job the cogitative power is supposed to do in order to make cognition possible.

Consider the following suggestion in reference to the animal estimative power:
Q55. It is interesting to note that in postulating the existence of this faculty, medieval writers clearly recognized that the potential for harm or good was not explicit in the perception itself. Instead the perception had to be interpreted in terms of previously experienced or innate knowledge. (Kemp, 1990, p. 58)

Information about danger or usefulness of an object is not contained in the object's accidental features cognized by the external senses. If it was, then not only sheep, but also cougars would perceive wolves as something dangerous and something to be fled. And so, this suggests that some innate internal mechanisms are necessary for the proper functioning of perception in animals. It has to be innate for a given species that certain phantasms produce certain behaviors. The vis estimativa certainly does not provide a sheep with an innate concept WOLF, nor does it tell her what a wolf is. It does impose on it, however, the capacity to take the data provided by the external senses, put it together in the right way in order to be able to determinately and accurately apprehend a particular concrete object and to cognize it as a 'wolf,' which, for sheep, stands for 'dangerous' and 'to be fled.' What is innate here is the structure or the functioning of the estimative power (or of animal instinct), which determines the appropriate discriminatory capacities that an animal possesses and the order of the series of actions that it undertakes that are required for survival. The estimative power has an intermediary role: it takes in sensory data, and on this basis it produces some specific kinds of behavior. A similar idea, I think, could be applied to Aquinas's doctrine of the vis cogitativa. Because the cogitative power is an instinct, it seems plausible that its
functioning is characterized by a similar sort of innateness as it is the case with the animal estimative power.

It is the fact that the cogitative power is a *sensory* power which guarantees the connection between our conceptual knowledge and information contained in our sense organs. This is what preserves Aquinas's cognitive empiricism. The capacity of the cogitative power to categorize, on the other hand, cannot belong to the level of the senses. Rather, it has to be explained by presupposing an innate structure of human cognition in which the intellect's concepts are available to the cogitative power. The latter, then, should be understood as an innate capacity to perceive the world as containing primary substances, and to assign these primary substances into various broader classes. Below (in 3.3) I suggest that the cogitative power is in fact an innate mechanism that cognizes things in the world through the transcendental concepts of the intellect, and that applies the intellect's intelligible species to the phantasms that it produces based on the data provided by the lower internal senses. In order to develop this kind of an account of the collaboration between the cogitative power and the intellect in which Aquinas's cognitive empiricism is reconciled with the kind of nativism that posits certain innate cognitive mechanisms, we need to present Aquinas's doctrine of transcendentals, and to explain how transcendentals can be used by the cogitative power (cf. below, 3.3 and 5.4.4).

Before we do that, let us notice that what the above discussion indicates is that for Aquinas, there are two levels of conceptualization in the human cognition: one

---

86 This capacity is similar to an innate ability that we have to perceive the world three-dimensionally.
purely intellectual (although the purity of the human intellectual cognition would be somewhat 'stained' by the fact that the human intellect never cognizes without turning to phantasms), the other also perceptual with the cogitative power as the main player. The way in which adult human beings perceive the world usually does involve conceptualization, but this does not imply that the intellect is directly involved. The cogitative power, Aquinas suggests, is sufficient for us to perceive primary substances, and even to see them as belonging to different categories. It is likely that the estimative power in non-rational animals also endows them with a capacity for some kind of conceptualization. In this case it is the natural instinct that determines an animal's capacity to categorize the objects that it encounters into different classes.

(iii) Humans vs. Animals

On Aquinas's account, there is no difference between humans and non-rational animals with respect to what we might call purely physiological sensation: in both cases it consists in the sensory organs registering various kinds of sensory stimuli while the stimulation actually occurs. As concerns the internal senses, although Aquinas ascribes them also to various non-rational animals, he believes that their functioning is different for humans and for animals. Aquinas explains:

Q56. with regard to sensible forms, there is no difference between a human being and other animals. For they receive a similar impression from external sensible things. But there is a difference with regard to the intentions under discussion. For other animals perceive these intentions only through a kind of natural instinct, whereas a human being also makes a kind of comparison. And so that which in other animals is called the natural estimative power is called the cogitative power in a human being. (ST I 78, 4)
What distinguishes the working of the internal senses in humans is the tight connection (and the resemblance) of the cogitative power to the intellect. Because Aquinas believes that the estimative power, which has no connection to the intellect, does not cognize individuals 'as existing under a common nature,' it may seem that cognition of individual substances is also an exclusive feature of beings endowed with the cogitative power. Consider, however, the following passage:

Q57. instinct is not aware of an individual thing as in a common nature, but only in so far as this individual thing is the term or principle of some action or passion. Thus a sheep knows this particular lamb, not as this lamb, but simply as something to be suckled; and it knows this grass just in so far as this grass is its food. Hence, other individual things which have no relation to its own actions or passions it does not apprehend at all by natural instinct. (InDA II 13, 398)

What the above passage suggests is that according to Aquinas the natural instinct in non-rational animals is sufficient to enable them to somehow cognize things as individual objects instead of, say, Humean bundles of scattered perceptions. As Aquinas explains, in non-rational animals,

Q58. the natural estimative power brings about the apprehension of an individual intention. It is in virtue of this that a sheep, through hearing or sight, recognizes its offspring or anything of that sort. (InDA II 13, 397)

And indeed, the wolf affects the sheep as an individual thing of a certain kind, and this is not just in virtue of its color or shape. What kinds of things the animal will react to, and in what way it will behave, is determined by the animal estimative power. Aquinas's point in the above passage is that for animals it is the possibility of action or
being acted upon that constitute "the limit of estimative knowledge" (Klubertanz, 1952, p. 202). A sheep recognizes a particular object as its offspring because, by natural instinct, it classifies it "as something to be suckled." A non-rational animal, as Aquinas explains (cf. *InDA* II 13), cognizes an object as it is a principle (or source, or an end) of an action or passion. On the other hand, "objects with respect to which there is no action or being acted on for the animal, are not known by it as individual" (Klubertanz, 1952, p. 202). If an object has no reference to action, that is, if the animal cannot intentionally act towards it and if it does not affect the animal beyond its external sensory faculties, then, properly speaking, the animal cannot cognize it (not by instinct in any case). That is why animals will not have such concepts as JUSTICE, ATOM or SYLLABLE.  

This suggests that for Aquinas (higher) animals are like humans in that they possess at least some kinds of concepts. Consider the following currently debated positions with respect to concept possession:

A. For a *concept Cartesian*, a creature can be said to possess the concept DOG just in case it can think about dogs, or have DOG-thoughts. For concept Cartesians, "it's not what you know (-how or -that) that determines what concepts you have; it's what you are able to think about. To have the concept DOG is to be able to think about dogs as such; and conversely, to be able to think about dogs as such is to have the concept DOG" (Fodor, 2004, pp. 30–31).

---

87 Wolves affect sheep, and wolves are made of atoms, so clearly atoms do affect sheep. But while it makes sense that the sheep could acquire the concept WOLF, she will not acquire the concept ATOM. Atoms do affect animals, but, properly speaking, they neither affect their sensory organs, nor their internal senses. And so, there is no consciousness of atoms as atoms. *Atoms as atoms* don't do anything to the organism's sensory faculties. They do affect the senses only because they constitute sensible qualities.
B. On the other hand, according to the minimal version of the pragmatist account, a creature can be said to possess the concept DOG if it can discriminate things that are dogs from things that are not. More generally, on this view, concept possession is some sort of dispositional condition; it is some sort of 'knowing how.' On a more robust version of the pragmatist view, that is, on the cognitivist account, in order to be ascribed the possession of the concept DOG, a creature must know what conditions must be satisfied in order for the concept to apply correctly to a given thing (cf. Fodor, 1998a, p. 124). Possessing the concept DOG is a matter of being able to 'know-how' to do something, and in particular of being able to draw certain inferences or to make certain perceptual judgments.88

While, for a Cartesian, concepts are thought-parts, for a pragmatist, they are mental abilities. On the pragmatist view, it is the behavior of a creature that provides evidence for whether or not it possesses a given concept. This—assuming that we take pragmatism in its minimal version—seems to be exactly the way in which, according to Aquinas, we can determine whether non-rational animals possess concepts. Animals (at least some kinds of higher animals) could be said to possess concepts understood as acquired mental dispositions that are expressed in behavior. Perhaps we could say that animals are like young kids: even though neither are able to explain what conditions must be met by a thing to fall under a concept, both must have some kind of access to

88 A cognitivist might consider concepts to be meanings of words, mental objects, or some other sort of abstract entities. Here the proof of the possession of a concept would most likely involve the use of language: to possess a concept, a creature would have to be able to formulate justified judgments involving that concept, or even to provide justification for judgments involving that concept. Requirements of this kind would only allow adult and otherwise 'normal' human beings to count as concept users. Still, it seems impossible for people to draw inferences or to make judgments without thinking, and so, without satisfying the Cartesian requirements for concept possession.
concepts because they do react to things as to *things of a certain kind*, that is, to things as falling under a certain concept.

For a Cartesian, as we said, a creature can be said to possess the concept of a given thing just in case it can *think* about that thing, or, in other words, if it possesses the appropriate thought-part. It is not clear whether non-rational animals, for Aquinas, can be said to satisfy the Cartesian requirements for concept possession. Aquinas would say, for instance, that both humans and animals possess the concept of 'something unpleasant.' The possession of such a concept can be expressed in behavior: in both an animal and a human the instinctive reaction to what is dangerous or unpleasant is to try to avoid it. It seems, then, that we need to rely on the pragmatist requirements for concept possession in order to figure out whether animals possess a given concept.

Now, when a human person sees something as falling under two concepts, e.g. under both the concepts 'unpleasant' and 'beneficial,' she can decide to suspend her instinctive behavior. If she thinks that the benefits would be large enough to make it worthy to suffer some pain, she will voluntarily accept what is unpleasant. This is because of a more general capacity of human beings to change their environment, to create new experiences and take an active part in the acquisition and even creation of concepts. Human intelligence allows them to control the association of concepts.

As concerns a non-rational animal say, a dog, it cannot *decide* to change her reactions with respect to various concepts. Animals have no control over what concepts they possess. It is rather their environment, or their experiences, that determine what
concepts the animal will associate, and what concepts it will keep in its memory. Of course, a dog could be conditioned, by the environment or by a human person, to change her behavior and accept something 'unpleasant' because of something 'beneficial.' However, even though the animal might differentially respond to different stimuli, we would not say that it has a grasp of concepts in question, because the animal is not in control with respect to what concepts it acquires and how the concepts that it has are associated.

The minimal pragmatist requirements for concept possession make it easier to determine whether a given creature possesses a given concept. It could be, however, that satisfying the pragmatist requirements already presupposes the capacity to think about referents of a given concept. If that's the case, then the problem with the Cartesian requirements of concept possession only indicates our own limitations in what we can figure out about animals. We must rely on how animals behave in order to ascribe to them any kind of beliefs or concepts.⁸⁹

Overall, the similarities between humans and animals that we have discussed confirm the intuition that cognition in general is, to a great extent, instinct-like, and, therefore, that in both animals and humans it should be interpreted as something innate.

---

⁸⁹ Observation of the behavior is not the only way in which we can decide whether a human being possesses a concept. On the contrary, what might be called the most genuinely human feature is the highly abstract kind of thought that is used, e.g., in pure logic and mathematics, or in building abstract physical theories and conceiving the idea of God. It may be that this kind of thinking could not be translated into any sort of behavior, or that it could only be expressed in language.
CHAPTER III

INNATENESS IN AQUINAS: THE LEVEL OF THE INTELLECT

Before we explain in more detail how the FTF explains the second level of innateness we need to provide a more thorough analysis of what concepts really are for Aquinas. I start with a presentation of the most common way in which Aquinas scholars interpret his views on concepts. We will see that this most common reading is unsatisfactory. Then I present what I consider a more plausible interpretation of Aquinas’s views on concepts.

3.1 Aquinas on concepts: Intelligible species vs. mental word

3.1.1 Concepts in Aquinas’s theory of cognition

Concepts, according to Aquinas, are immaterial forms of things abstracted from phantasms and existing in the intellect; they are similitudes (or intellectual likenesses) of material things (cf. e.g., Stump, 2003, p. 262). They are means of cognition: they are not what (quod) we cognize, but that by means of which (id quo) we cognize (even though they can become, secondarily, objects of cognition). Aquinas explains:
Q59. the intelligible species is related to the intellect as that by which it understands (…).
That by which the sight sees is the likeness of the visible thing; and the likeness of the thing understood, that is, the intelligible species, is the form by which the intellect understands.

(…) intelligible species is that which is understood secondarily; but that which is primarily understood is the object, of which the species is the likeness. (…) the soul knows external things by means of its intelligible species. (ST I 85, 2) [species intelligibilis se habet ad intellectum ut quo intelligit intellectus (…). similitudo rei visibilis est secundum quam visus videt; et similitudo rei intellectae, quae est species intelligibilis, est forma secundum quam intellectus intelligit. (…) Et sic species intellectiva secundario est id quod intelligitur. Sed id quod intelligitur primo, est res cuius species intelligibilis est similitudo. (…) anima per species intelligibiles cognoscat res quae sunt extra animam.]

All of these descriptions do not yet provide a satisfactory definition of concepts. As we recall from chapter I, for Aquinas, there are two kinds of vehicles of cognition at the level of the intellect: intelligible species and mental words—both sometimes translated as concepts. Since Aquinas believes that the intellectual cognition always requires a return to phantasms, before we try to figure out how intelligible species and mental words should be understood, let us first consider in what way both of them differ from phantasms.

(i) Phantasms vs. Concepts

In On Spiritual Creatures Aquinas explains the difference between vehicles of cognition at the intellectual and sensory levels as follows:

Q60. a species which is in the [internal sense of] imagination is of the same genus as a species which is in a [external] sense, because both are individual and material. But a species which is in an intellect belongs to another genus, because it is universal. (DSC 10, ad17)

In the Summa Theologiae, he offers a similar explanation:
Q61. the sensible image in sense is the likeness of only one individual thing, and can give the knowledge of only one individual. But the intelligible species of our intellect is the likeness of the thing as regards its specific nature, which is participable by infinite particulars; hence our intellect by the intelligible species of man in a certain way knows infinite men; not however as distinguished from each other, but as communicating in the nature of the species; and the reason is because the intelligible species of our intellect is the likeness of man not as to the individual principles, but as to the principles of the species. (ST I 14, 12)

Phantasms, which are the proper objects of the internal senses, turn out to be very different from concepts. Concepts are immaterial and universal. Phantasms, as Aquinas explains in the *Summa Theologiae*, are individual in the sense of being likenesses of particular material things (*ST* I 85, 1, obj. 1: *phantasmata sunt similitudines particularibus*). Like sensible species that are modifications of sensory organs, phantasms are also supposed to be material. This is because, Aquinas explains, "a phantasm must be connected with the extended and time, from the very fact that it is a likeness of a singular thing which is 'here and now'" (*InDMR* 2, 314).

Notice that there are two different ways in which materiality of vehicles of cognition can be understood. Typically, materiality is understood as an ontological property. It is said that for Aquinas both sensible species and phantasms are material in the sense that they are realized physically in the human body. This feature is then contrasted with concepts' immateriality—concepts, on this reading, are not physically realized in any way. There is no need, however, to interpret Aquinas in this implausible way. A better solution is to suggest that the materiality of phantasms, like their individuality, refers to the kind of representation that each of these vehicles of cognition provides. In the following passage from *De Veritate* Aquinas makes a distinction between (a) the ontological status of vehicles of cognition, that is, what kind of thing
they are in the cognizer's mind, and (b) their function, that is, what they represent, and in what their relationship to the object cognized they remain.

Q62. All cognition follows some form which is the principle of cognition in the knower. Such a form can be considered under two aspects: either with relation to the being it has in the knower, or in the reference it has to the thing it represents. (DV 10, 4)

(a) So, first of all, both phantasms and concepts (as well as sensible species) can be considered as existing, that is, in their ontological status as they actually occur in the thinker's mind. (In fact, this is the only way in which phantasms and concepts actually exist. There is no such a thing as the phantasm 'cat.' Rather, there are only individual occurrences of a 'cat' phantasm, each occurring in an individual cognizer, and caused by an individual cat, or by some particular stimuli produced in the cognizer's imagination.) With respect to their ontological status, phantasms and concepts are similar (cf. Pini, 1999, p. 31): as a phantasm or a concept occurs in the cognizer's mind, it is a particular entity of a certain kind, it is something different from the object it represents; it is an individually existing thing, a mental item, an accident (an individual modification) of the human mind. In modern terms, we'd say that each is somehow physically realized in the brain. In this sense, from the point of view of their ontological status, both can be called material. (Aquinas does not have to claim that concepts by means of which we think are some mysterious immaterial entities not having anything to do with the human brain.)
(b) The function of vehicles of cognition is to make cognition possible. For this they have to be somehow connected to things in the world; they have to represent things in the world. Both phantasms and concepts are mental representations. But they represent different things, and in a different way. The difference seems to be straightforward: through concepts we cognize the constitutive features or the essence of a thing, what makes the thing be what it is. Through phantasms we access only accidental features of the object cognized. Phantasms represent properties of objects existing in the external world; concepts represent their universal essences, that is, things that we do not encounter in the external world. (As we said in the beginning, things in the external world are always particular, and always composites of form and matter.)

This characterization of phantasms and concepts leads to the following problems. If all that phantasms represent are accidental forms of things, while the intellect's proper objects represent essences then (i) there is still no solution to the Transduction Problem (cf. (ii) in Introduction, problem P3); there is no explanation of the passage from sensible species to concepts (introducing phantasms does not help because both sensible species and phantasms represent only accidental forms) and (ii) it seems that human cognizers have no cognitive access to things in the world, to individual substances, which are the basic ontological category for Aquinas (problem P5 in (ii) in Introduction). In addition, on this account to have a concept means to know the essence

---

90 According to Claude Panaccio (2001), the understanding of concepts as mental representations has an immediate consequence for their ontological nature: they must be something intermediary between the object cognized and the cognizing subject. This, according to Panaccio, implies that Aquinas cannot be a direct realist. This view is not shared by other Aquinas scholars, for instance, by Gilson (1961) and Kretzmann (1991). Cf. above 1.2.2 (iii).
of something. However, we don't always manage to figure out an essence of things that
we encounter in the world, but we can still think about them or cognize them
intellectually (cf. below, 3.3). So, the idea that the proper objects of the intellect are
essences seems to require too much from intellectual cognition.

As a solution to these problems I propose an interpretation of Aquinas's account
of cognition that will include three claims:

1. intelligible species and mental words should be understood as, respectively, concepts
   as thought-parts, and concepts as conceptions;
2. there are different kinds of phantasms, depending on which one of the internal senses
   is involved in their production;
3. the highest kind of phantasm which is operated on by the cogitative power is in some
   sense equivalent to intelligible species.

(ii) The Hierarchy of Phantasms

The idea that on Aquinas's account there are different kinds of phantasms is confirmed,
for instance, in the SCG 2, 73 where Aquinas says that when phantasms "are understood
potentially," which means, when they are considered from the point of view of the
internal senses (where they are only potentially intelligible), they "are of various kinds."
Because different internal senses have importantly different roles, it makes sense that
they would not all operate on the same kind of proper objects.

(A) Phantasms of the lowest kind are the products of the common sense. The common
sense (cf. above, 1.1.2 and 2.3.2) is an integrative internal faculty which takes the various
sensible forms registered by each of the external senses (the three-dimensional colored expanse, the single specific odor, the meowing noise, etc.), and unifies them by putting them into a percept which represents these features as existing in a determinate physical configuration. The phantasm produced by the common sense is similar to sensible species in that it occurs only while the physical stimuli are actually present. It's different in that it is a representation of sensory data already integrated into a determinate physical configuration.

(B) The output of the sensus communis further stimulates the internal senses of the higher level, imagination and the estimative and cogitative powers. Phantasms of the second kind are produced by the internal sense of imagination or, in Aristotle's case, by phantasia (see above, 2.3.4). Imagination, as we have seen, is responsible for memory; it is a place where sensible species are stored (before this stage in the process of cognition they are present only while the stimulus is affecting the cognitive faculties of the cognizing subject). And so, phantasms at this level are like leftover impressions from sensory images of properties of particular things. They are traces of the unified percept that remain in the subject when the stimulus is already gone. This is what Aquinas talks about in his commentary on Aristotle's On Memory:

Q63. Something is known by a secondary movement which remains after the first change worked in the senses by the sensible object. This movement remains at times even after the sensible objects are gone and pertains to the phantasm. (InDMR 2, 319)
On the understanding according to which the role of phantasia is to transcend physiological limitations of external senses (cf. 2.3.4), the external senses work only when the object of cognition is present; that is, they register information only when the stimulus is actually affecting the cognizing subject; they register the imprints of the perceptible form only at a present moment. The internal sense of phantasia, on the other hand, produces phantasms that represent a panoramic view of a given situation. These phantasms are what allow us to understand the whole sentence instead of a string of separate sound, etc. The phantasms produced by phantasia include more than what is given by the external senses.

Phantasms of this kind, we could say in more contemporary terms, are involved in what Dretske calls non-epistemic seeing. The sensible species produced by the external senses are not sufficient to allow the subject to see, for instance, a Boeing 747. When only the external senses are at work the subject only registers discrete sensory features. Phantasia allows us to see a Boeing 747, but still without recognizing it as one. If no further internal sense is involved, seeing that something is a Boeing 747 is not yet possible. The latter kind of seeing, seeing that, is more conceptual (even though it is still based on sensory evidence). It involves some kind of beliefs or knowledge. The kind of seeing that phantasia facilitates is merely sensory.

Phantasms that are produced by the internal sense of estimative power allow the subject to recognize various things that the external senses by themselves cannot perceive. They organize and present the objects from the external environment under
the categories of attraction, repulsion, or indifference. They are, in fact, interpretations of what is sensed. The interpretation is determined by the needs and desires of the organism according to its biological heritage and individual experience.

Phantasms from levels B and C are different from sensible species in that they can in-form the cognizing subject even when the sensory stimuli are not there anymore, and in that they contain more information than what's given in the output of the external senses. They are similar to sensible species (and different from the intelligible species) in that they only represent different aspects of individual accidental features of the object cognized that were contained in the percept produced by the common sense.

(D) The most important internal sense for human beings (and unique to them) is the cogitative power. This sense faculty is very different from other internal senses in that it collaborates with the intellect. As we have seen, it is responsible for two things: cognizing individual substances, and cognizing them as being of a kind. The phantasms produced by the cogitative power, therefore, represent the individual substance that is the subject of the accidental features represented in a percept; they represent this substance as an individual substance of a kind.

My suggestion is that this highest kind of phantasm that the cogitative power produces, and which is unique to human beings, is in a sense equivalent to the intellect's intelligible species. The intelligible species, on the account I propose, are applied to the phantasm produced by the cogitative power as symbols. An intelligible species is a symbol of a corresponding phantasm produced by the cogitative power.
All phantasms are representations of incidental sensibles—of things that the external senses cannot cognize on their own. Phantasms of the highest kind are representations of a specific kind of incidental sensibles, of particular material objects as substances of a kind. The cogitative power does not represent accidents, but it cognizes individual substances. All phantasms, also those produced by the internal sense of cogitative power, contain more information than what is contained in sensory impressions (cf. Pasnau, 2002, pp. 271 and 277).

Because of its collaboration with the intellect, the cogitative power possesses some kind of classificatory power. The cogitative power sorts out individual objects by means of some kind of the intellect's innate programming. It is, however, not yet capable of a full blown conceptualization. (It is not yet seeing that something is a Boeing 747.) It allows us to see a Boeing 747, to recognize another thing as the same as the other Boeing 747, and to distinguish it from other kinds of objects. But representing a Boeing 747 as an individual substance of a kind does not entail seeing or recognizing it as a Boeing 747.

Based on the output produced at a lower level of cognition the cogitative power produces a phantasm that is a representation of an individual object, and a representation of an object of a certain kind. This automatically provides space for the intellect's intelligible species—for symbols that will stand for any substance of the same kind. The cogitative power in a way sends the following information to the subject of cognition: "this thing in front of you is an (...)" where the gap (...) waits for a concept that will be provided by the intellect. (What we might wonder here is whether the phantasm produced by the cogitative power clearly indicates a particular kind that the thing it
represents instantiates. After all, a Boeing 747 is an instance of many different kinds of things: it is an airplane, a passenger plane, a Boeing product, etc. We would not find an answer to this question in Aquinas. What seems likely is that for an ordinary person the phantasm produced by the cogitative power represents a medium-sized easily recognizable kind, such as an airplane (or a dog), and not a sub- or superordinate kind, such as a Boeing product (or a Chihuahua), or a flying object (or a mammal.)

### 3.1.2 Intelligible species vs. mental word

We said that for Aquinas both intelligible species and mental words are sometimes translated as concepts. In what follows I'm going to argue that *intelligible species* and mental words are not the same thing. On the contrary, the difference between them will turn out to be crucial. It is only intelligible species that should be understood as concepts. Intelligible species, I'm going to claim, are the basic tools of the intellect—they are the tools by means of which the intellect *thinks*. Mental words, on the other hand, are the ultimate goal and the final products of successful intellectual cognition. The intellect's intelligible species are applied to the highest kind of phantasm, the phantasm produced by the cogitative power. They constitute a link between the internal senses and the intellect. Mental word, on the other hand, is the highest kind of vehicle of cognition produced solely by the intellect; it is a representation of the thing's essence. It is something that never shows up at the level of the internal senses.
For Aquinas, the intellect is opposed to the senses in that it cognizes essences of things, while the senses cognize accidental aspects of reality. Essence is the proper object of the intellect: essence of the object cognized is what is represented by the intellect's concepts. This, as we said, sounds problematic once we recall that essences (in the mind) are always universal and things that we encounter in the world are always individual. If essence is a proper object of the intellect, and so, the only kind of object that the intellect can operate on, then it seems that the intellect cannot cognize objects in the world because the intellect on its own has no access to individuals. This sounds a lot like Platonism all over again.

Aquinas defines an essence as the nature of a thing as it is given in a definition (DEEE 1, 4). This suggests that intellectual cognition of a cat would always consist in having a grasp of what really is the essence of cats. Indeed, on the common interpretation of Aquinas, when you intellectually cognize a cat, this means that you must have a mental grasp of something which is distinct from perceptible qualities, and which seems very rich in content. This content is supposed to inform you of what it really means to be a cat. It is supposed to represent the classical definition of 'cat' in terms of its genus proximum and differentia specifica. In addition, it is commonly agreed that for Aquinas it is impossible to have this conceptual grasp of an essence of cats without also having sensory experience of cats. So, human cognition, as Aquinas supposedly presents it, is very mysterious indeed: the senses provide information: pink, large, moving, etc. Then, somehow on this basis, the intellect thinks: 'homo sapiens.' And in addition, the former is a necessary condition for the latter.
Here is what I think could help us get rid of at least some of the mystery. My suggestion is that when Aquinas says that the proper object of the intellect is essence, there are actually two cases that should be distinguished. These two cases correspond to the two ways in which Aquinas understands concepts: as species intelligibilis and as verbum mentis. What we need to consider now is the sense in which both intelligible species and mental words can be said to stand for essences of things. Let us look at the following passage:

Q64. the intellect informed by the species of the object, by understanding produces in itself a kind of intention of the object understood, which intention reflects the nature of that object and is expressed in the definition thereof. (SCG 1, 53)

The passage suggests that the intelligible species is a vehicle of cognition operative during an earlier stage of the intellect’s activity. On the other hand, the mental word, referred to as 'a kind of intention,' is the end-result of the intellective cognition. A mental word is something produced by the intellect in the process of understanding; it represents the nature of the object cognized as it is expressed in a definition. So we could say that intelligible species and mental words as two distinct kinds of vehicles of cognition correspond to the two operations at the level of intellectual cognition: thinking and understanding. Aquinas explains:

Q65. The action of understanding is not exercised without something being conceived in the mind of the one who understands, and this is called the word: since before a concept of some kind is fixed in the mind we are not said to understand but to think about a thing in order to understand it. (DP 9, 9)
Understanding is the most sophisticated kind of human cognition. It is on this level that we get to know the essence of the object cognized. As Aquinas explains,

Q66. whenever we understand, by the very fact of understanding there proceeds something within us, which is a conception of the object understood, a conception issuing from our intellectual power and proceeding from our knowledge of that object. This conception is signified by the spoken word; and it is called the word of the heart signified by the word of the voice. (ST I 27, 1)

Q67. when I wish to conceive the nature of a stone, I must get there through reasoning. So it is for all things that are understood by us. (…) As long, then, as the intellect is tossed this way and that by so reasoning, its formation is not complete—until it has completely conceived of the thing’s very nature, and then for the first time it has the character of a complete word. (InIoan I, 1, 26)

For Aquinas, to understand, or to truly know what something is is to be able to define it (cf. In I Sent 37, 3, 3). To give a definition of something means to provide its genus proximum and differentia specifica (cf. InBoeth 6, 4, etc), to say what genus it belongs to, and in what it differs from other members of the genus. Understanding, therefore, is something that we achieve by reasoning; it consists in arriving at the 'thing's very nature,' it involves knowledge about the object. If a mental word is the end-product of this process, then it is plausible to interpret it as something close to what today we would call a conception, that is, a theory-like mental representation which "denotes a type of belief system" (Davis, 2003, p. 416).

On the other hand, before the intellect produces this rich-in-content mental word, there is a prior stage of the intellect's operation, thinking, in which, as Aquinas says in the passage from SCG 1, 53, the intellect is 'informed by the species of the object.' Thinking requires much less than understanding. If the intellect is informed by the
species, then there certainly has to be some connection with the object. The content of intelligible species (i.e., what it represents) is still determined by the object. But this content does not have to be the object's essence. What thinking requires is only mental signs—certain tools that allow the intellect to refer to the object; tools that the intellect uses for thinking. Intelligible species, therefore, can plausibly be understood as concepts in the sense of thought-parts. They are mental signs that may stand for or point at the quiddity of a thing\(^{91}\) of a certain kind; we don't need to know, however, what this quiddity is.

Consider also the following passage:

Q68. [Aristotle] mentions another kind of unity, namely that which comes of a thing being one in kind, though made up of discontinuous parts, e.g. the unity of a man, or a house, or even of an army. (...) the soul (...) understands it by what is undivided in the soul and in an indivisible point of time (...). And though division may be contained in such unities, the divided parts are not understood—so far as the object and time of the understanding are concerned—as divided, but as united; for even though there be an actual division into parts, the species itself, as such, is indivisible; and this it is that is indivisibly understood. But if the parts are understood separately—e.g. the flesh and bones and so forth—the whole is not understood in an undivided time. (InDA III, 11, 755)

The way I propose to understand this passage is as showing that, for Aquinas, a concept, as a thought-part, does not express the definition of the thing cognized. Rather, a (simple) concept as a thought-part does not have a structure.\(^{92}\) As Aquinas puts it, "the species itself, as such, is indivisible." And also: "The soul (...) understands it by what is undivided in the soul and in an indivisible point of time" (InDA III, 11, 755). A concept

---

\(^{91}\) This is because, as Aquinas puts it, "a thing is intelligible only through its definition and essence" (DEEE 1, 4). The concept DOG as a thought-part represents an object which is an instance of the universal essence dogness.

\(^{92}\) That is, a simple concept in its content has no structure. There can be, of course, complex concepts. And also, the object that a concept stands for may have a structure.
as a thought-part is merely a *sign* pointing in two directions: on the one hand, it points towards an individual object that belongs to its extension; on the other hand, it points towards the object's essence, that is, towards the definition of the *kind* to which the thing belongs. Possession of such a sign does not require an understanding of the essence that it is a sign of. On the other hand, we can take mental words to be *conceptions*. A mental word expresses the thing's essence or a definition: to possess such a conception means to understand what a given thing really is. There can be no question of the mental word's being indivisible (except perhaps in case of a mental word representing something absolutely simple).

So, my conclusion here is that in the most basic sense, concepts for Aquinas are thought-parts. Although they still point towards universal essences, they do not express them.

The interpretation of intelligible *species* as concepts in the sense of thought-parts and of mental words as conceptions indicates that the suggestion that the proper object of the intellect are universal essences *does not* mean that for Aquinas 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 the human cognitive activities to understand the essence of the thing cognized, and so to acquire a mental word. However, it is enough for us form concepts as thought-parts in order to be able to think about things whose essences we do not grasp.

I don't claim that Aquinas himself intended this interpretation. Perhaps the way I use the terms he's been using is not exactly what he'd want. Still, it seems that my
interpretation makes Aquinas's account of concepts more intelligible for the contemporary reader; it shows that his account is plausible and interesting, and that it is not inconsistent in any startling way with Aquinas's other views on cognition.

3.2 The Transcendentals

Until now we've been trying to make sense of Aquinas's claim to the effect that *essences* are the proper objects of the intellect and to reconcile it with the idea that the intellect thinks by means of concepts. In the *Summa Theologiae*, we see Aquinas also saying that "being is the first intelligible as sound is the first audible" (*ST* I 5, 2: *ens* (...) *est primum intelligibile, sicut sonus est primum audibile*). Then, in *De Veritate*, he claims that "that which the intellect first conceives (...) and to which it reduces all its concepts, is being" (*DV* 1, 1). As we mentioned above (in 1.1.3 (iv)), these two claims, namely, that the intellect cognizes essences, and that being is the first object of the intellect seem to contradict each other because on Aquinas's account, essences in the mind are always universal (in the sense that numerically distinct things can have the same essence), and transcendental Being (*Ens*) always refers to what is individual (everything is a being, or is intelligible, in virtue of its own unique act of transcendental existence). If essences are the only kind of thing that the intellect can operate on, then it doesn't seem plausible to say that Being is the first intelligible. In order to make more sense of the claim that being

---

93 See also *In I Sent* 8, 1, 3; *DV* 21, 1; 21, 4, ad4; *DEEE* prologue; *DP* 9, 7, ad15; *ST* I 5, 2; 11, 2, ad4; *ST* I-II 55, 4, ad1; 94, 2.
is the first object of the intellect we need to clarify two things: (1) what 'being' actually means for Aquinas and (2) in what sense it is supposed to be \textit{first}. The second question will be considered below in 3.2.2 and 3.2.3. In 3.3.2 we will show how the two claims, i.e., that essences are the proper objects of the intellect and that what the intellect first cognizes is Being, can be reconciled.

\section*{3.2.1 Aquinas's conception of being}

On the most common understanding, 'being' is synonymous with 'physical object' or with 'individual substance.' This is what our everyday use of the word would suggest: we use the word 'being' in reference to physical objects. And this is not surprising. We are material beings, and, like animals, we are in our cognition dependent on sensory experience which gives us access to individual features of individual substances. Knowledge by sense is temporally prior to knowledge by intellect; and knowledge about sensible things is temporally prior to knowledge about non-sensible things (cf. Bobik, 1965, p. 3). Aquinas would not want to deny any of this. He agrees that the most obvious objects that the intellect cognizes are individual physical things. He believes that because of its cooperation with the internal senses which are programmed to detect and register

\footnote{The explanation of the concept \textit{BEING} that I will provide here is far from exhaustive. My intention, however, is only to provide what I consider a plausible explanation of Aquinas's understanding of the concept \textit{BEING}.}

\footnote{I will use the capitalized \textit{Being} for what the concept \textit{BEING} refers to; this kind of \textit{Being} does not really exist; it is not something that we could find in the world. When I use the word 'being' without any capitalization, I refer to particular existing things and their properties — to anything that has some kind of existence.}
individual substances in the world, the intellect's most natural operation is, precisely, cognition of the essences of individual substances. (As we have seen above, it is 'being' in exactly this sense—as an individual object—that is the proper object of the cogitative power.)

The understanding of being as an individual object—I'll call it Being-0—is one among several that we can find in Aquinas's writings. Being understood as an individual object, such as a cat, or a stone, is the most natural, but not the most fundamental way to understand this concept. When Aquinas talks about 'being' as the first intelligible, as he does in the quoted passage from ST I 5, 2, he has something else in mind. Before I identify this more fundamental sense of being, however, it will be useful to introduce another distinction made by Aquinas himself. In De Ente et Essentia Aquinas explains:

Q69. the term 'a being' in itself has two meanings. Taken one way it is divided by the ten categories; taken in the other way it signifies the truth of propositions. The difference between the two is that in the second sense anything can be called a being if an affirmative proposition can be formed about it, even though it is nothing positive in reality. In this way privations and negations are called beings, for we say that affirmation is opposed to negation, and that blindness is in the eye. But in the first way nothing can be called a being unless it is something positive in reality. In the first sense, then, blindness and the like are not beings.

The term 'an essence' is not derived from this second meaning of 'a being,' for in this sense some things are called beings that do not have an essence, as is clear in the case of privations. Rather, 'an essence' is derived from 'a being' in the first meaning of the term. As the Commentator says, a being in the first sense of the term is that which signifies the essence of a thing. And because, as we have said, 'a being' in this sense is divided by the ten categories,

96 Anthony Kenny distinguishes twelve ways in which the term 'being' appears in Aquinas's writings (cf. Kenny, 2002, pp. 189–192). Being as B-0, i.e., as 'individual physical thing,' corresponds to Kenny's number 3 (substantial being) and to some extent to his number 4 (accidental being).
97 This 'something else' is closest to Kenny's number 5, that is, to what he calls 'common being,' something that all things that exist have in common, or a very thin and universal predicate (Kenny, 2002, p. 190; cf. also SCG I 26, 241).
essence must mean something common to all the natures through which different beings are placed in different genera and species, as for example humanity is the essence of man, and so with regard to other things. (*DEEE* 1, 2-3)

So, to the above mentioned B-0, i.e., being as an individual substance, we can now add two further senses of being:

Def. B-1 — whatever belongs to one of the ten categories is a being; anything that is a substance (a thing of a certain kind), and also any real or positive property of a thing (any of the nine kinds of accidents) is a being. The common and crucial feature of all beings in this B-1 sense is that they all have an essence; each of them constitutes "something positive in reality" (*DEEE* 1, 2). (Aquinas makes a similar point in *DV* 1, 1 where he states that "every reality is essentially a being"). B-0 is included in B-1 (B-0 is a subset of B-1).

Def. B-2 — anything that you can form an affirmative proposition about is a being (some of these don't have an essence — negations or privations, e.g., blindness).

B-0 and B-1 are included in B-2.

According to B-1, 'being' is not synonymous with 'individual substance,' because substance is only the first one of the ten categories. 'Being,' on the other hand, applies to any thing which falls under any of the ten categories. As Kenny explains it,

Q70. We may say (...) of Thomas Aquinas that he was a human being, and that he was fat, clever, holier than Abelard; that he lived in Paris, in the thirteenth century, that he sat when

---

98 According to Kenny, this condition has to be qualified: it would be more correct to say that in the B-2 sense of 'being,' all beings are such that 'true' affirmative propositions could be formed about them (Kenny, 2002, p. 4).

99 According to Aquinas, even though negations and privations (and also things that are evil) can be defined, they do not have an essence, they are not ontologically real. When they are defined, this can be done only by means of negative terms, indicating a lack of some positive (ontologically real) property. For instance, Joe's hatred for Jack, on Aquinas's account, is a being in B-2 sense (it can be defined), but it does not have a transcendental essence. For this reason it would not count as being in what I call B-B sense below.
lecturing, wore the Dominican habit, wrote eight million words, and was eventually poisoned by Charles of Anjoy. The predicates we use in saying these things belong, Aristotle would say, in different categories: they belong in the categories of, respectively, substance, quantity, quality, relation, place, time, posture, vesture, action, and passion. (Kenny, 2002, pp. 2–3)

On this B-1 understanding of the term, not only Aquinas as a person is a being, but also Aquinas’s intelligence, and even his being holier than Abelard. This is because, Kenny continues,

Q71. When a predicate in a particular category is actually true of something, then, according to Aquinas, there exists in the world an entity corresponding to the predicate. (Kenny, 2002, p. 3)

Cognitive interaction with things that fall under one of the ten categories leads to acquisition of substance- or accident-concepts (universal concepts). It does not lead to the acquisition of the transcendental concept BEING (and especially not to the acquisition of the mental word BEING; more on that below). In its most natural operations the intellect does not cognize 'being-as-such,' but it cognizes individual cats, chairs, or patches of green. The objects that the intellect cognizes are first of all individual physical things (at least this is where the intellect starts; this is because we are, like non-rational animals, dependent on the senses in our cognition). This kind of cognition manifests itself in the formation of concepts of these individual things and their qualities. The intellect, however, can also arrive at the concept of BEING-AS-SUCH.
Being as B-2 has an even wider scope than B-1. On this understanding, anything that you can think or talk about is a being. This suggests that to be intelligible and to be a being are one and the same thing. As Aertsen explains,

Q72. Being is that on the basis of which things are capable of being known by an intellect, it is the prerequisite condition for every intelligible object, for something is intelligible insofar as it has being. (Aertsen, 1996, p. 80)

This makes sense because if something is cognizable (if you can think about it) at all, then indeed there has to be something intelligible in it—otherwise the intellect would have nothing to hook itself to. But this understanding of being also leads to a problem. The intellect can think of such things as Santa Claus. After all, we understand what Santa Claus is, and we can think and make true statements about it. So it would seem that Santa Claus must be a being as well, and —according to B-2—that there must be something intelligible in it. The problem here is that it is not quite clear what this 'it' is supposed to refer to.

It seems that we have two options here: one would be to say that it is only our thought about Santa Claus that is a being. For Aquinas, my thought about Santa Claus certainly is a being (in the B-1 sense of the term): thoughts understood as particular mental states are accidents of the cognizing subject. They belong to the category quality. When the prophet thinks ‘quid est Israhel, quid est quod in terra es inimicorum, inveterasti in terra aliena?’ (Baruch 3, 10-11), he formulates his thought and the words by means of which he expresses it so that they have a certain determinate sense. The prophet is wondering about the fate of the nation of Israelites, and not of Philistines; and he wants
to know the reason of their misery, as opposed to anything else he might think about them. The concrete thought entertained by the prophet, and the words that he could use to express it, are certainly something different from nothingness (the only 'thing' that is not a being is exactly a non-being, and so, nothingness). The prophet understands them in one specific way, and not another. They differ from other thoughts that the prophet could entertain at other times and from his other mental states, etc. The prophet's thought, therefore, certainly qualifies as a 'being.'

But what about the object of the prophet's thought? Is 'the misery of Israelites in the enemy land' a being? And in general, is every object of thought a being as well? Is Santa Claus a being? It seems that Aquinas's answer to these questions would also be 'yes.' Like the thought itself, the object of thought—the intentional object to which the thought about Santa Claus refers, is not a nothing. It is a determinate object of thought, distinct from other things one could think about, and it makes a certain concrete sense (if it didn't make sense, it could not be an object of thought). So it seems that it would be correct to call it a being. Let us see how Aquinas explains the sense in which privations and negations are also 'beings.' He says in De Veritate:

Q73. Being is, in some way, predicated of non-being in so far as non-being is apprehended by the intellect. For, as the Philosopher says, the negation or the privation of being may, in a sense, be called being. Avicenna supports this by pointing out that one can form propositions only of beings, for that about which a proposition is formed must be apprehended by the intellect. Consequently, it is clear that everything true is being in some way. (DV 1, 1, ad7)

100 It qualifies as a being in the B-1 sense and also as a being in the B-B sense—the transcendental sense of being that will be discussed below in 3.2.2. Because both B-1 and B-B are subsets of B-2, anything that is a B-1 or a B-B is also a B-2.

101 It qualifies as a being in the B-B sense and as a being in the B-2 sense.
A similar explanation that applies in the case of negations and privations could also be used with respect to different kinds of intentional objects. However, also within intentional objects Aquinas would most likely distinguish two classes: those intentional objects that represent ontologically real things, properties or states of affairs (i.e., intentional objects that have an essence), and those that refer to things, properties or states of affairs that are not ontologically real. (Another kind of objects that we could worry about is possible objects of thought. Are these also beings? Or is it the case that only actual objects of thought or those actually thought about in the past count as beings?) In any case, the ontology of intentional objects and of negations and privations goes beyond the scope of the present discussion. For now, let us agree that 'being,' for Aquinas, applies to the following kinds of things:

(i) physical objects, independent, substantial beings (and so, we can say that John exists) and their properties, non-independent accidental beings (John's aquiline nose exists); the latter (accidents) include John's relationships to other things and also his mental states;
(ii) material beings (this here horse exists) and immaterial beings (my thought of a triangle exists; my love for person XY exists);
(iii) natural beings (this here oak exists) and artificial beings (this locomotive exists);
(iv) necessary beings (the intentional being-principle of contradiction exists) and non-necessary beings (this here hairdo exists) (cf. Krapiec, 1991, p. 91).

If 'being' can mean so many different things, what sense do we ascribe to it when we call it the first object of the intellect?
3.2.2 Being and the other transcendentals

As it turns out, in addition to the different understandings of 'being' mentioned above, there is also, for Aquinas, what can be considered the most fundamental sense of 'being,' being-as-such, that refers to the transcendental property or transcendental concept BEING—I'm going to call it B-B. It is Being in this most fundamental sense that is the first object of the intellect that Aquinas talks about in ST I 5, 2, DV 1, 1, and other places.

(i) Transcendentals—a general description

The theory of transcendentals lies at the very foundation of Aquinas's philosophy. As Breton puts it, "La doctrine des transcendentaux est une des clés de la métaphysique thomiste (…); classique et pourtant méconnue" (Breton, 1963, p. 45). Aertsen admits that one of the goals of his book is "to show that the theory of the transcendentals is foundational for Thomas's thought" (Aertsen, 1996, p. 71; cf. also Aertsen, 1996, p. 19, where he makes an even more daring statement; transcendentals, he suggests, are the foundation of the whole of medieval philosophy: "Over time, my conviction has grown

102 On this understanding, 'being' is not only something that has some reality in itself, or something that has an essence, but also something that allows us to arrive at the transcendental concept BEING. It has to be somehow independent of the ten categories because they can only lead us to universal concepts. B-B is a subset of B-2 (the latter has the broadest scopes; it includes all possible senses of Being).

103 Unless indicated otherwise, in the remaining part of this chapter the term 'Being' always refers to being in the B-B, that is, in the transcendental sense of the word.

104 In Aquinas's writings, there is no separate treatise devoted to transcendentals in general. In his early work, De Ente et Essentia, Aquinas provides an extensive treatment of the first two transcendentals, Being and Thing. The most important sources for Aquinas's views on transcendentals are: In I Sent 8, 3, and also De Veritate 1, 1 and De Veritate 21, 1.
that medieval philosophy can be regarded as a way of transcendental thought, as a *scientia transcendens*.

Notwithstanding its fundamental importance for Aquinas’s philosophy, the theory of transcendentals is often ignored by Aquinas scholars: we won’t find separate chapters on transcendentals in books on Aquinas’s philosophy. As Aertsen puts it, the

> Q74. importance for the transcendental way of thought is often neglected or underestimated. Although libraries have been written on his thought, and although we possess various studies of some aspects of his doctrine of the transcendentals, it is striking that a comprehensive study of Aquinas’s doctrine is lacking. (Aertsen, 1996, p. 22)

Aertsen doesn’t try to suggest any rationale of this state of affairs. One possible explanation could be found in Krapiec’s suggestion:

> Q75. Although cognition based on the use of transcendental predicates is the most general, it is also the least precise, the least informative. Consequently, a metaphysics based on this type of cognition would be the most general but also the least useful kind of knowledge; it would be nothing more than a collection of current slogans. (Krapiec, 1991, p. 11)

Transcendental concepts are the most general concepts that there are. They apply to everything and anything that we encounter in the world. For this reason, however, their informational content is minimal. So it may be that the doctrine of the transcendentals is ignored simply because it seems that it just doesn’t add anything to the scope of our knowledge. Another reason for why transcendentals are so often overlooked is perhaps that it may seem that they are a part of Aquinas’s metaphysics; it is not obvious that they are also crucial for his account of cognition.
Interestingly, a similar 'silent' treatment by Aquinas scholars can be observed with what concerns Aquinas's account of the cogitative power and of internal senses in general (cf. Lisska, 2000, and Peghaire, 1942-3). The theory of the internal senses is often ignored, probably because it is considered too unclear and muddy, to be considered a crucial element in Aquinas's account of cognition.

So, we have two elements of Aquinas's philosophy that generally do not receive much attention. And even when they are not ignored, they are never to my knowledge considered together. My suggestion is that putting together these two elements of Aquinas's philosophy not only provides solutions to the problem of transduction mentioned above, but also is exactly what would allow us to present Aquinas's account of cognition as being on a par with theories of cognition currently discussed.

(ii) How many Transcendentals?

Different philosophers provide different lists of transcendentals. Some of the lists overlap, some don't. Most philosophers mention the following three transcendentals: Being, Thing, and Unity. Aquinas also adds: Something, Good, Truth and Beauty. The

---

105 Most often those scholars, who talk about one of the elements, ignore the other: in his *Metaphysics*, Krapiec spends a lot of time discussing Aquinas's theory of transcendentals, but doesn't even mention the internal senses. Kenny (1993) talks about the internal senses, but doesn't mention transcendentals; Wippel (2000) talks about transcendentals, but not about internal senses, or the cogitative power in particular. There is no mention of transcendentals in Stump's *Aquinas*. Pasnau's 2002 mentions them once and very briefly (on p. 326) – as something perhaps obvious but not requiring further explanation; certainly not as something that in Aquinas's view counts as fundamental for human cognition.

106 Aertsen discusses an interesting explanation of the origin of transcendentals: all of them are really names applying to God: God is who is; God is one; only God is good; Jesus calls himself truth, etc. (cf. Aertsen,
list of transcendentals that Aquinas provides is not generated randomly. On the contrary, there is a specific order in which he thinks transcendentals can be derived. Transcendentals, for Aquinas, express different modes of being. The list of transcendentals is determined by two elements: the features of being itself and the way in which our intellect can access it. As Krapiec puts it, the list of the transcendentals is generated "not arbitrarily, but in an order designated by the structure of reality" (Krapiec, 1991, p. 107).

The intellect can focus on being per se, or on being in relation to another (ens in alio). This is how transcendentals are divided into two groups: absolute and relational, depending on whether "the expressed mode of being pertains (...) to every being in itself (in se) or in relation to something else (in ordine ad aliud)" (Aertsen, 1996, p. 98). Being, Thing, and Unity are absolute transcendentals because they focus on being itself; they apply to every existing thing independently of whether anything else exists; no relation of the given being to another is needed (cf. Breton, 1963, pp. 48–50). Transcendentals Something, Good, True, and Beautiful are all relational. They refer to being in relation to the mind (or the will) of another being, or another cognitive subject (a person) (cf. Krapiec, 1991, p. 104).

---

1996, p. 21). Since God is 'first,' and in some way the most fundamental for the whole of the universe, it makes sense that what is the most fundamental for things in the world gets its name from features of God. (This explains why Aristotle didn't list 'good' and 'truth' among transcendentals: Aristotle didn't have the concept of creation, or of everything that exists as flowing from the First and One cause that is at the same time good and true.)

107 See a very good explanation of three slightly different ways in which the transcendentals are derived in Aquinas's three basic texts in Aertsen, 1996, pp. 98–102.

108 I will not talk about the transcendental Beauty in this dissertation. There is no agreement among Aquinas scholars with respect to the issue of whether Beauty is a separate transcendental, or whether it is reducible to Good, Truth and Unity. Cf. Aertsen's schema of the systematization of the transcendentals in Aertsen, 1996, pp. 100–101.
Within the *absolute* transcendentals, Being and Things are *positive*: they affirm something of every being *per se*. Unity or One is negative: it denies something of every being in itself: it says of every being that it is undivided, that it is not a non-being. The relational transcendentals also can be further divided into negative and positive. There are two kinds of relation in which one being can stand to another: negative in the case of the relation of two things which are distinct from one another, and positive—when there is some kind of correspondence between two things. The negative relation is expressed by the transcendental Something (*Aliquid*) which indicates that to be a being is to be 'divided from others' (*DV* 1, 1). The positive relation is the relation of correspondence or conformity of one being to another. According to Aquinas, "there is something which is such that it agrees with every being" (*DV* 1, 1). Everything that exists, for Aquinas, is such that it is in a relationship—at least potentially—to a cognizer's soul (or to the Cognizer's, i.e., God's, soul). And since the soul is endowed with both cognitive and appetitive faculties, it is related to every being in two ways: there is a correspondence between being and the appetitive part of the soul which is expressed by transcendental Good, and a correspondence between being and the intellect, expressed by transcendental Truth.

In what follows I only provide a summarized description of the first six transcendentals.\(^ {109} \) Transcendentals Being, Thing, Unity and Something, which are

---

\(^ {109} \) This part of the dissertation is an attempt to present in as clear a way as possible Aquinas's account of transcendentals as it is explained mostly by Krapiec. I am aware that a much more extensive study would be needed to give justice to the issue of transcendentals.
instances of first principles, that is, among others, of laws of identity, non-contradiction, and excluded middle, are the most important for my purposes in this dissertation.

(iii) Being and Thing

Being (Esse) is the first and the most important transcendental. Anything positive in reality, anything real is a being. As Aquinas states in De Ente et Essentia, "nothing can be called a being unless it is something positive in reality" (DEEE 1, 2: "non potest dici ens nisi quod aliquid in re ponit").

The fundamental act of every thing is its existence in the transcendental sense, the fact of its being (ontologically) real. Also, our cognition is inseparably tied to transcendental existence. According to Aquinas, "without being, nothing can be apprehended by the intellect" (In I Sent 8, 1, 3: "primum enim quod cadit in imaginatione intellectus est ens, sine quo nihil potest apprehendi ab intellectu"). The intellect can only cognize something that exists (in the transcendental sense of the word), something that is in act (cf. Krapiec, 1991, p. 107). We cannot encounter or talk of something that is not a Being. So, Being is both the first and necessary condition for anything to exist, and for anything to be cognized. As Krapiec puts it, "the existential dimension of being makes possible the emergence of cognition and constitutes the most original 'object' of cognitive apprehension for the human being" (Krapiec, 1991, p. 109).

110 "That which the intellect first conceives, as best known, and in which it resolves all it conceptions, is being (ens)" DV 1, 1 (see also In I Sent 8, 1, 3; DV 21, 1; 21, 4, ad4; DEEE prologue; DP 9, 7, ad15; ST 1 5, 2; 11, 2, ad4; ST I-II 55, 4, ad1; 94, 2 and more). In these passages Aquinas talks about Being as one of the transcendentals, and so, about Being in the B-B sense. This kind of Being is what the mind primarily detects (cf. below, 3.2.4 (iv) — at least the human mind in its present state where it is (typically) dependent on the senses in its cognition. The mind can also be considered a mechanism detecting being in the B-2 sense, but only in a secondary sense.
It is precisely because transcendental Being is the most fundamental feature of anything that is, that it is very difficult to express in words what exactly this transcendental is. We might try to define a Being as 'that which exists,' 'what is,' 'what is in actuality,' 'what participates in the act of being,' or 'what has being.' All of these definitions, however, already direct us beyond the transcendental Being, with its focus on nothing but the existence or actuality of some object, and indicate the object's other, essential, aspect. In the phrase 'that which exists,' 'which exists' stands for the thing's existence, and 'that which' stands for the thing's essence. So it turns out that we cannot define Being without immediately referring to the second transcendental: Thing. Still, Ens on its own does not yet signify some thing that is; rather, it "abstracts from the subject of being (subiectum essendi) and signifies the act as such indeterminately" (Aertsen, 1996, p. 187). At the same time, Ens does not signify something abstract, it does not signify existence in general. This can be explained if we contrast the verb Esse with the participle Ens. Think of it as analogous to the difference between 'to run' and 'runner' (esp. since Ens is also translated as a noun 'being'). The former signifies abstractly, the latter signifies a concretum. Ens is like the (act of) running of a particular runner in that it "signifies the act as concretized in a subject which has it" (Aertsen, 1996, p. 187); it always refers to an individual thing-in-actuality.

Notice that Ens is not limited to physical existence, nor to actually and physically existing individual substances. Indeed, for Aquinas, substance is the most important meaning of Being: it is something that is in the proper sense. Transcendental Being, however, goes well beyond it. We can think about Napoleon even though he is dead, or
about Pegasus even though it never enjoyed physical existence. To exist, for Aquinas, is
to be in actuality in the sense of being intelligible.

The transcendental Thing (Res, Essentia) emphasizes the fact that anything that
exists must be a thing of a certain kind; it must have an essence; there must be a set of
features in terms of which this thing can be described, and that make it the kind of thing
it is. We are built in such a way that once we are aware of something as existing, we
immediately know that it will be of a kind, that it has an essence. We have to think this
way—that's how we are constructed and so determined to cognize the world.

Our most fundamental cognitive contact with reality always has this dual aspect:
we notice both that there is or exists something, but we also immediately or
simultaneously know that for anything to exist it has to be some thing, it has to have an
essence. And so, even though Being (and the existential aspect of things) is more
fundamental than Thing, still, "no human cognition can detach itself from the essential
aspect, the content aspect" (Krapiec, 1991, p. 110). (It's interesting how rarely we
consciously think about the existential aspect of things. The fact that things exist, that
there is something rather than nothing, is the most amazing thing about the world. For
Aquinas, actually, the existence of anything at all—or at least the existence of any
contingent thing—is already a proof of God's existence. God is a continuous per se cause
of anything that exists—He sustains things in existence.)

---

111 As Aertsen points out, some later medieval traditions actually deny that res is a separate transcendental
That Being is more fundamental than Thing can be seen for instance in the fact that the latter is actually derived from the former: Essence is derived from Being (DEEE 1, 2-3). If we want to be the most precise with the sense of the transcendental 'Being,' then it should be taken not as a noun, but as a participle (cf. Tavuzzi, 1987, p. 572). A 'being' is something that exists, something that is in the act of existing, and, in this sense, is doing something. Being as a noun, on the other hand (and here is where Thing appears) is derived from the verb 'to be.' In Aquinas's own words, "thing differs from being because being gets its name from to be, but thing expresses the quiddity or essence of the being" (DV 1, 1). This, as Krapiec notices, "points to the primacy of the existential dimension of being both in our cognition as well as in being itself" (Krapiec, 1991, p. 109).

Notwithstanding the fact that Being is the first and most fundamental transcendental, we talk of things more commonly than of beings. Aquinas offers an explanation of this phenomenon. On his view, human cognition could never proceed without the essential aspect of the object. Without the essential aspect, we would not be able to formulate any judgment about the world. This is because (transcendental) existence itself cannot be grasped by concepts (understood as mental words). Existence of a thing is something that can be experienced, and noticed, but it is very difficult to communicate this experience to another. Whenever we try to transmit our experiences to others, we do it by means of conventional signs (concepts as thought-parts) that immediately take us back to what a thing is, and which go beyond the fact

---

112 Aquinas explicitly claims that (transcendental) existence cannot be grasped by concepts. This claim is not plausible, however, if we take concepts to be thought-parts. It makes sense, however, if we think of concepts as mental words (or conceptions).
that it just is (cf. Krapiec, 1991, p. 108). Every judgment contains concepts that stand for properties of things, for what something is, and not for the mere fact that it is.

(We can now better see the difference between the transcendental Being, B-B, and the different senses of being mentioned above. B-0, B-1 and B-2 all refer to the 'whatness' of a thing, to what something is. Being as the first object of the intellect, on the other hand, refers to be-ing—to the very act of existence, to the fact that something is, the very facticity of a thing.)

It is important to bear in mind that although Essence and Thing do not have exactly the same meaning (we say, e.g., that a dog is a Thing, and that it has an essence; cf. Krapiec, 1991, pp. 114, 107), still, they are sometimes also used as equivalent (one reason for this is that only something which has an Essence can be a Thing). So, for instance, the title of one of Aquinas's early treatises, for which his main reference is Avicenna's *Metaphysics*, is *De Ente et Essentia* (*DEEE*)—*On Being and Essence*. In the Prologue of *DEEE* Aquinas says: "the first conceptions of the intellect are (as Avicenna says) 'a being' and 'an essence.'" Interestingly, however, Avicenna does not explicitly mention Essence, but rather he talks about *Thing* (see Aertsen, 1996, p. 160). Aquinas gives the name 'Essence' to what Avicenna calls 'Thing.' (The fact that Essence and Thing can be used interchangeably is crucial for my purposes. I talk about it in detail in 3.3.2.)

---

113 In Aquinas's prologue to *DEEE* we read: "A slight initial error eventually grows to vast proportions, according to the Philosopher. Now the first conceptions of the intellect are (as Avicenna says) 'a being' and 'an essence.' If, then, we are to avoid mistakes through ignorance of these, we must begin exploring their difficulty by stating what is meant by saying 'a being' and 'an essence,' how they are found in different things, and how they are related to the logical notions of genus, species, and difference."
Once we reflect upon the first two transcendentals, we notice that together they express one of the first principles, that is, the Law of Identity, \( p \equiv p \). Defining Being in terms of Thing and vice versa is an instance of this law. We define Being as 'that which exists.' This sounds like a tautology, like saying nothing more that Being is being. The phrase, as Krapiec explains, actually says more because the identity between definiendum and definiens is only relative. The definiens expresses both transcendentals at the same time: as we said, 'which exists' stands for the thing's existence; it designates the act of existence of something "without any more specific cognitive qualifications—something that appears as a result of an initial intellectual contact with reality" (Krapiec, 1991, p. 111). 'That which' stands for the thing's essence—for "a determinate content of that which exists" (Krapiec, 1991, p. 110). So, the Law of Identity here says that Be-ing, that is, some being-in-actuality, equals Being of a Certain Kind; Being equals Being that has an Essence. Even though Being and Thing always refer to the same object and designate the same individual concrete reality, they allow us to look at that object from different perspectives. "In all creatures, as Boethius has pointed out, 'to be is other than that which is'" (DV 1, 1, ttc3).

---

114 For Aquinas, the first four transcendentals (Being, Thing, One and Something) are expressions of the way in which the intellect sees reality: the intellect must cognize the world through two most fundamental first principles, the Law of Identity and the Law of Non-Contradiction. First principles govern not only the functioning of the human mind, but are also the basic objective laws that govern the universe.

115 Krapiec calls the relationship between Ens and Essence 'relative identity.' What he means by it is merely, or so it seems, that in the definition of Being as 'that which exists' the definiens and the definiendum have different meanings but the same referent.
(iv) One

The next transcendental, One (Unum; Unity, Unit), is a negative absolute transcendental which says that every being is in itself — that is, as a being — undivided: it is made one by its one act of transcendental existence. As Aquinas puts it, "There is (...) a negation consequent upon every being considered absolutely: its undividedness, and this is expressed by one. For the one is simply undivided being" (DV 1, 1). In the Summa Theologiae Aquinas explains: "The one does not add a reality to being, but only the negation of division" (ST I 11, 1). In his commentary on Peter Lombard's Sentences, Aquinas defines One as follows: "Unum est ens quod non dividitur" (In I Sent 24, 1, 3, ad3; cf. also DP 7, 7).

The transcendental One emphasizes the fact that for something to be a being, it has to possess unity; it has to be some one thing, and it has to be one and the same thing. This way the transcendental One emphasizes a different aspect of Being. As it was the case with Being and Thing, transcendental One also expresses the Law of Identity, but differently formulated. It says that 'every being is a single thing,' or 'every being is one and the same thing' (or, more as a tautology: Being equals One Being). The law of identity thus formulated contains all three absolute transcendentals. (Transcendental One not only stands for non-division, but it also points back to the transcendental Thing
and Being. Similarly, for Aquinas, all subsequent transcendentals contain the previous ones.\(^{116}\)

In addition to the Law of Identity, the transcendental One as a metaphysical property of things can also be considered an instance of the Law of Non-Contradiction \(\neg(p \& \neg p)\) (cf. Krapiec, 1991, p. 118). The Law of Non-Contradiction says that nothing that exists can be a non-being, or that being cannot at the same time be a non-being (cf. Krapiec, 1991, pp. 116, 103). When we analyze any existing thing, we notice that there is no internal division in it: no being has in itself anything that would not be this very being. As Aristotle puts it, "it will not be possible for the same thing to be and not to be" (Metaph IV 4, 1006b19). We can see, therefore, that what transcendental One expresses is "the impossibility of joining in one judgment an assertion and denial of the same object" (Krapiec, 1991, p. 119).

Now, it is an obvious fact that most things (that we encounter in the world) have parts. If to be a Being is to be One and undivided, then are all the things which have parts not Beings? Aquinas considers this issue in various places. In the Summa Theologiae he says:

\[Q76.\text{ What is compound, has not being whilst its parts are divided, but after they make up and compose it. Hence it is manifest that the being of anything consists in undivision. And hence it is that everything guards its unity as it guards its existence (esse) (ST I 11, 1).}\]

Then, in his commentary on Aristotle's De Anima, Aquinas explains:

\(^{116}\text{Cf. Metaph IV 1003b22-25: "Being and unity are the same (...), for one man and a man are the same thing and existent man and a man are the same thing." In fact, for Aquinas, "each of the previously formed transcendental concepts serves as a basis for subsequent concepts" (Krapiec, 1991, p. 104).}\]
Q77. [Aristotle] mentions another kind of unity, namely that which comes of a thing being one in kind, though made up of discontinuous parts, e.g. the unity of a man, or a house, or even of an army. This is a specific, not a quantitative indivision; and the soul, he says, understands it by what is undivided in the soul and in an indivisible point of time (...). And though division may be contained in such unities, the divided parts are not understood—so far as the object and time of the understanding are concerned—as divided, but as united; for even though there be an actual division into parts, the species itself, as such, is indivisible; and this it is that is indivisibly understood. But if the parts are understood separately—e.g. the flesh and bones and so forth—the whole is not understood in an undivided time. (InDA III 11, 755)

If some thing, even a thing that has parts, is an existing being at all, it is because all its parts are unified by its one act of existence. It is the thing's Esse which makes it one being, and which makes it a being in the first place. In fact, we can talk about division and a thing having parts only with respect to the thing's nature; existence itself is always one and only one act. As Aquinas explains,

Q78. what exists might have added to itself something extraneous to its nature, [but] existence itself cannot receive anything that is not pertinent to it. (DQO 4, 600)

(v) Something

The transcendental Something (Aliquid, Separatness) emphasizes the fact that every given being is distinct from all other things: not only is it not divided in itself, as was indicated by the transcendental One; in addition, it is also distinct from every other concrete being. As Aquinas puts it, "Being is called something insofar as it is divided from others" (DV 1, 1).

The transcendental Something is an expression of the Law of the Excluded Middle (p v ~p), which says that "there is nothing between being and non-being; there is
no middle" (Krapiec, 1991, p. 135). We have seen that the transcendental One, or Unity, divides Being from non-being. The transcendental Something does the same thing, but in a different way. It separates any given existing thing, which in itself is a unity, from other beings which, in relation to the being in question, are exactly non-beings. (They are non-beings in a relative way: in relation to the originally given thing, all other things are not-this-being, and so, they are non-beings; cf. Krapiec, 1991, p. 134.)

The dividedness from other things is also a necessary condition for cognition. We could neither perceive nor intellectually cognize this one particular thing if it didn't differ from other things. Also, our intellect can think of this one particular thing only because it sees it as different from other things. In this way, the transcendental Something indicates the pluralistic structure of the world: if something can be a being (or an object of cognition) only if it is distinct from other beings, then there must be those other beings in the first place.

As we said, For Aquinas transcendental concepts are interchangeable. This seems quite intuitive when applied to the first four transcendentals. Anything that exists is a being (esse), but something can be a being only if it is a thing of a certain kind, and if it has a determinate set of properties, or a determinate essence (res). Also, for something to be a thing of a certain kind means to be an instance of that kind (unum). Finally, to be a thing of a certain kind means not to be a thing of a different kind (aliquid). The transcendentals Being, Thing, One and Something all have the same scope. They are the same with respect to reference and only differ in sense—they emphasize different aspects of what it means to be a being.
The remaining transcendental s, Good and Truth, may seem more odd philosophically. They express the idea that anything that exists is, respectively, good and true. First of all, however, it seems rather commonsensical to think that some things are not good. Furthermore, it is not clear what it could mean to say that all things are true. We are used to applying truth only to propositions or statements, and not to things. The last two transcendental s, therefore, require a further explanation.

(vi) Truth (*Verum*)

As Aquinas explains in *De Veritate*, "truth expresses the correspondence of being to the knowing power" (*DV* 1, 1). In fact, Aquinas distinguishes two kinds of truth. He says: "truth can be considered as existing in the thought or in the thing itself" (*ST* I 39, 8). On the one hand, and more commonly, we talk about logical truth—the kind of truth which applies to sentences, propositions or statements. This kind of truth is the goal of our cognition. It consists in a correspondence between our intellect and things in the world. In order to verify whether our cognition is true, we need to check if it matches what really is the case in the world. Our beliefs, or our statements, are true if they correspond to how things are in reality. The second kind of truth that Aquinas talks about is *ontological* truth or truth of things themselves. This kind of transcendental truth applies to anything that exists, to every being. This truth, for Aquinas, also consists in a correspondence between reality and the intellect, but it is a correspondence that has an opposite direction of fit. According to Aquinas, everything that exists is created by God, and God creates things in the world by *thinking them*. Things in the world, therefore, are
true in the sense that they correspond to what is in God's mind. In Aquinas's words, true is "that which imitates the divine exemplar or is in a relationship to a cognitive power" (In I Sent 8, 3). Any being is true to the extent to which it agrees with the divine idea of it. Ontological truth, therefore, expresses the fact that being for Aquinas is dependent on the intellect of the Creator. This is what is supposed to show that reality makes sense, it is cognizable, intelligible. It can be cognized by intellects, human intellect included, but its intelligibility is not the product of the human intellect. The ontological truth of things shows that intelligibility of the world is not relative to the human cognizer. Rather, as Aertsen puts it, "truth resides solely in the thing itself, expressing its intelligibility, and is thus identical with being" (Aertsen, 1996, p. 246).

(vii) Good (Bonum)

Goodness, for Aquinas, is a transcendental property that refers to a thing's desirability. A thing is desirable as a thing of a certain kind to the extent to which it is actualized as a thing of this kind. Aquinas explains:

Q79. The essence of goodness consists in this, that it is in some way desirable. (...) a thing is desirable only in so far as it is perfect, for all desire their own perfection. But everything is perfect so far as it is actual. Therefore it is clear that a thing is perfect so far as it exists, for it is existence that makes all things actual. (ST I 5, 1)

What is desired must at least be perceived as desirable, and if something is desirable, it is only because it is good in the transcendental sense of goodness (see for example SCG I 37 and III 3). A dog is good in this sense not when it is a dog that can fly or that can explain Gödel's theorem. Rather, a good dog is a dog that actualizes its specific
potentialities (potentialities that differentiate its species from another) and that realizes its dog-nature to the full. In general, any existing thing is desirable as a thing of its kind and hence good of that kind to the extent to which it is actualized and in being (ST I 5, 1). 'Being' and 'goodness' have the same referent: the actualization of specifying potentialities.

After the brief characterization of the transcendentals as they are understood by Aquinas, we need to see in what way they can appear to be the tools of the intellect and the innate mechanisms necessary for cognition that would help overcome all the problems mentioned above, and how this would make Aquinas's account of cognition overall more plausible.

3.2.3 What are transcendentals?

What Aquinas refers to as 'transcendentals,' can be viewed in two ways:

(i) Ontological perspective on transcendentals

From the ontological point of view, transcendentals are the fundamental properties of every existing thing. Every thing that exists has three kinds of features: individual, universal (or generic), and transcendental. Individual features of a thing are what make it unique and distinct from everything else that exists. Universal features of a thing make it belong to a given class of things; it is because of its universal properties that a
thing is placed in a given species or genus. Universal properties are properties that apply to one class of things and differentiate this class from another. Finally, transcendental features apply to any individual existing thing just in virtue of its being a being. Transcendental features go beyond—transcend—all the boundaries between different kinds of things, and apply to anything that exists.\footnote{The terminology used by medieval scholars is confusing. Because transcendentals are properties of every existing thing, then they seem to qualify as universal properties. Their universality, however, is so to speak 'more universal' than that of universal properties as they are usually understood. While a universal property 'dogness' can be instantiated only by dogs, and the property 'chair' only by chairs, a transcendental property is instantiated by everything that exists or that is intelligible: it transcends different categories of things.}

As Aertsen puts it, transcendentals "precede everything and are firsts in things" (Aertsen, 1996, p. 19). Transcendental properties apply to every thing not because of what it is, of what kind of thing it is, but because it just is in the first place. Anything that exists is, \textit{first of all}, a being, it is something (and so, it is different from nothing), it is a thing of a kind, it is identical with itself, it is distinct from other things, etc. All other features of things are in a way secondary—an object may possess a non-transcendental feature only assuming that it possesses all of the transcendental properties.

Transcendentals, therefore, are the most general and universally applying properties of things. They apply of necessity to every being \textit{qua} being. As Aquinas explains in \textit{De Veritate}, "the mode [of being] they express is one that is common, and consequent upon every being" (DV 1, 1). If transcendentals don't apply to something, then this 'something' is \textit{not} a being (for this reason Aquinas would say that in an important sense evil is not real—it is not a being\footnote{Evil is not a being in the B-0 or in B-B sense; it can be called a being in the B-2 sense. See above, 3.2.1.}).
(ii) Epistemological perspective on transcendentals: universal vs. transcendental concepts

As we were talking about transcendental vs. universal properties of things, we can also distinguish between transcendental and universal concepts that correspond to these properties. The first feature in which transcendental concepts differ from universal concepts is their scope.

Q80. Universal terms have a strictly defined scope due to the univocal content of their designates, which form a determinate group of individuals clearly distinct from other groups. (…) Transcendental terms, on the other hand, represent an unlimited content, a content of infinite scope, for every object can be called 'good,' 'being,' 'one,' or 'true.' (Krapiec, 1991, pp. 10–11)

Universal concepts are limiting; when one universal concept applies to a thing, this means that another universal concept will not apply to it. If something is a cat, then it is not a dog; if it is (entirely) black, it cannot be (entirely) green. This is not so with transcendental concepts. On the contrary, transcendentals "go beyond genera and categories, but they do it only to reunite them. The transcendentals bring together what is separated by the categories" (transcendentals "dépassent les genres et les catégories, mais pour les relier. Le transcendental unit ce que la catégorie distingue," Breton, 1963, p. 49). Transcendental concepts apply to everything in reality. They indicate what we can say about any thing without applying any concrete property to it. We may have no idea what something is, we might be unable to ascribe any non-transcendental property
to it (or to use any of the individual or universal concepts in reference to that thing), but
still, we are able to say a lot about this thing just because it is.

Unlike universal concepts, transcendentals are interchangeable, that is, if one
applies to something, then all other transcendentals apply to it as well. Anything that is
a being, that exists, is also a thing, it is one thing of a certain kind, it is identical with
itself, is distinct from other things, and so forth. All transcendental concepts have exactly
the same scope, they are coextensive.

Although transcendental concepts are convertible and identical in their scope,
they are, as we have seen, not synonymous. As Aquinas explains in De Veritate, every
transcendental expresses "a mode of being that is not expressed by the name being itself"
(DV 1, 1, ad ttc1). All the subsequent transcendentals "explicate conceptually what being
is"; they "add something conceptually to being" (Aertsen, 1996, p. 97). As Aertsen
explains,

Q81. The general modes of being explicated by the other transcendentals manifest the
different 'faces' of being. They are an explication of being according to the aspect of its reality
(res), indivisibility (unum), division from others (aliquid), knowability (verum), and
appetibility (bonum). (Aertsen, 1996, p. 104)

On Aquinas's account, transcendental and universal concepts also differ in that the latter
apply univocally and the former—analogically. When we apply a univocal predicate to
two distinct things, we apply it in the same way to both of them. So, the predicate 'man'
applies in a univocal way to both Joe and Harry. Joe and Harry are 'the same' in every
respect which counts when we want to apply the predicate 'man' to them—even though
they differ with respect to their physical traits, personalities, etc. Transcendental concepts, on the other hand, only apply to things analogically: each thing is good, true, beautiful, and even, each thing is a being—to a different degree. As Aristotle says, "there are many senses in which a thing may be said to be" (Metaph IV 2, 1003a33). For Aquinas also each distinct being is a being in its own way. All transcendental properties are predicated of different things in different ways.

As we mentioned before, transcendental and universal concepts differ also with respect to how much they broaden the scope of our knowledge. Transcendental concepts are unlimited in their scope: they apply to anything and everything. But the kind of knowledge they provide is very limited; it is the kind of knowledge which doesn't tell us anything specific about a given particular thing, e.g., what this thing is, or what non-transcendental properties it has that make it be the kind of thing it is (cf. Krapiec, 1991, pp. 10–11). Transcendentals only indicate the most universal and fundamental laws of being in general.

Universal and transcendental concepts differ, finally, in how they get acquired (or at least this is the claim I want to make in this dissertation). According to Aquinas, there are two fundamental ways in which a subject can come to possess a cognitive feature: naturaliter and through acquisition. Naturaliter, he believes, is the way in which we come to possess the first principles. In his commentary on Aristotle's Metaphysics, Aquinas explains that the knowledge of the first principles

Q82. is not acquired by demonstration or by any similar method, but it comes in a sense by nature to the one having it inasmuch as it is naturally known and not acquired. For first
principles become known through the natural light of the agent intellect, and they are not
acquired by any process of reasoning but by having their terms become known. (InMetaph IV
6, 599)

He also explains:

Q83. Our intellect knows being naturally, and whatever is per se comprised under being as
such; and on this knowledge is based the knowledge of first principles, such as the
incompatibility of affirmation and negation, and the like. Consequently, these principles
alone are known naturally by our intellect. (SCG II 83, 1678) (Naturaliter igitur intelletus
noster cognoscit ens, et ea quae sunt per se entis inquantum huiusmodi; in qua conditione fundatur
primorum principiorum notitia, ut non esse simul affirmare et negare, et alia huiusmodi. Haece igitur
sola principia intellectus noster naturaliter cognoscit.)

In what concerns the process of acquisition, it can be further divided into:

A. 'normal' acquisition which has sensation as its starting point and its foundation;
B. the 'sophisticated' kind of acquisition (here more reflection and conscious effort
is needed to acquire a given cognitive feature).

The first way, A, is how most universal concepts (from all ten of the Aristotelian
categories) are acquired in 'normal circumstances' and in 'normal' human cognizers. This
is the most 'regular kind' of cognitive process which Aquinas also calls 'abstraction' (cf.
above, 1.2.3 (vi)). As he says in De Veritate, "in statu viae, in quo per species a rebus
abstractas intelligimus" (DV 10, 11, ad10), that is, "in this state of life we cognize by means
of species abstracted from things." We should remember, however, that, as we have
concluded in 1.2.3 (vi), abstraction understood according to its common interpretation,
that is, as a purely empiricist process of ignoring certain features and focusing on others,
does not work. The normal kind of concept acquisition cannot be a purely empiricist process; rather, it requires the involvement of innate mechanisms.)

The second way, B, is the way in which the most abstract theoretical concepts get acquired — concepts in such disciplines as theoretical physics or pure mathematics. Clearly, most people never acquire concepts of this kind.

How we will explain acquisition of transcendental concepts will depend on how we understand them. On the one hand, it is correct to understand transcendental concepts as all other kinds of concepts, that is, as thought-parts (cf. 3.1 above) which appear in language as names and which designate properties. The unique feature of transcendental concepts is that they can also be interpreted as cognitive mechanisms. (As we shall see, it is solely in this sense that they can be truly called 'the firsts' in the process of cognition.) Depending on which meaning of the transcendentals we have in mind, a different kind of acquisition will have to be considered.

When transcendental concepts are understood as thought-parts, we would say that a subject may come to possess them by means of what we called acquisition of a more sophisticated kind (B above). This, Aquinas explains, is because

Q84. Those things which are farthest removed from the senses are difficult for men to know; for sensory perception is common to all men since all human knowledge originates with this. But those things which are most universal are farthest removed from sensible things, because the senses have to do with singular things. Hence universals are the most difficult for men to know. (InMetaph I 2, 45)

---

119 We could acquire the concept EXISTENCE by abstraction, but not TRANSCENDENTAL EXISTENCE. Most people never think of transcendental properties of things.
What’s the most universal, Aquinas believes, is the most difficult for human beings to know. Transcendentals have an even greater scope than any of the universal concepts, and so they are the most difficult to acquire. In fact, it may be that transcendental concepts as thought-parts are never acquired by a given person—it may be that a person never explicitly thinks of the world as containing beings, unities, etc., in the 

transcendental sense of these terms 120—or they are acquired quite late in the cognitive development of a subject. This is because they are the most abstract, and the "farthest removed from the senses" (InMetaph I 2, 45); they are difficult for men to know, and for this reason they can only be acquired and operated on by wise people.

Transcendentals understood as mechanisms, on the other hand, are acquired naturaliter. We will try to clarify in what sense they can be understood as mechanisms in 3.3. Right now we want to explain what it is supposed to mean that transcendental concepts are acquired naturaliter. This is connected to Aquinas’s claim that transcendentals are first in cognition.

Actually, the idea that they are first seems to imply both: that it is a good idea to understand them as mechanisms, and that they get acquired naturaliter.

120 Not many people, besides professional philosophers ever acquire these concepts. The transcendental sense of the terms ‘being,’ ‘thing,’ ‘one,’ etc., is rather counterintuitive to our everyday thinking. One way to explain it is that "ens commune, ens inquantum ens, which is the direct subject matter of metaphysics (…) has a content which is not at all readily or immediately seized by the intellect. Rather, it has a content which, if it is to be seized at all by the mind, requires the performance of a highly sophisticated process of resolutive reasoning" (Tavuzzi, 1987, p. 559). Regular language users employ such words as ‘being,’ ‘thing,’ etc., but not with their transcendental sense. Rather, it is most common to reduce, for instance, the concept BEING to the concept of INDIVIDUAL PHYSICAL THING, and this is, obviously, different from the transcendental understanding of BEING.
3.2.4 Logical vs. Psychological priority

(i) Different kinds of priority

Aquinas holds that transcendentals precede and are a necessary condition for intellectual cognition in general; without them, no intellectual cognition is possible. In Aquinas's words, "without being, nothing can be apprehended by the intellect" (In I Sent 8, 1, 3). In this context transcendentals clearly could not be understood as concepts in the sense of thought-parts: as we just said above, transcendentals as thought-parts are either never acquired or acquired late in the cognitive development.

How transcendentals in their role of first and necessary condition for cognition should be understood, and how we come to possess them can be inferred from their relationship with first principles. As we said above, naturaliter is the way in which we come to possess first principles. My suggestion is that to say that we have first principles naturaliter it means to say that this is how we are built to think: we have to think by means of these principles. First principles, therefore, when they are considered a necessary condition for knowledge, should not be understood as propositions, but as mechanisms, as necessary operational structures of the human intellect.

According to Aquinas, our knowledge of first principles is derived from or based upon our knowledge of Being. And Being itself is also known naturaliter:

Q85. Our intellect knows being naturally, and whatever is per se comprised under being as such; and on this knowledge is based the knowledge of first principles, such as the incompatibility of affirmation and negation, and the like. Consequently, these principles alone are known naturally by our intellect. (SCG II 83, 1678) (Naturaliter igitur intellectus noster cognoscit ens, et ea quae sunt per se entis inquantum huiusmodi; in qua conditione fundatur

200
Being is a basis for first principles, and they are a necessary condition for thinking. First principles must be known if we are to understand or know anything at all, or even just to think. So, if we think, we use first principles, but we can only use them if we have them—that is, if they have been derived from Being, which we also have naturaliter. Since Being as a concept is acquired late, it follows that Being that we have naturaliter is not a concept (not a thought-part).

Both first principles and transcendentals that we have naturaliter are best understood, I suggest, as innate mechanisms or as tools that the intellect uses to think. As Aquinas puts it, "the agent intellect is like an artisan and the principles of demonstration are like tools" (QDA 4, ad6). Since, as we have seen (cf. above, 3.2.2), first principles, i.e., the Law of Identity, Non-Contradiction and Excluded Middle, etc., are the cognitive, epistemological expressions of the transcendentals as metaphysical properties of things, this means that when we talk about first principles and transcendentals in this context, we really have in mind one and the same thing: innate mechanisms of cognition. (I talk about it below, 3.3.1.)

At this point we need to distinguish between the ontological order of priority—the order of what is first, and the cognitive order of priority: the order of what we know first (cf. Aertsen, 1996, p. 161). The latter, first of all, should not be taken in the sense of the temporal order. When Aquinas talks about transcendentals as being first, he is not concerned with "the order of the temporal genesis within individual consciousnesses of
particular concepts” (Tavuzzi, 1987, p. 556). Temporal priority concerns the order of the
temporal acquisition of particular concepts in a subject’s mind (most likely our
temporally first concepts are concepts MOTHER, FOOD, etc.). Cognitive firstness of the
concept BEING (and of other transcendental concepts) does not mean that thoughts
containing the concept BEING are the first thoughts that we ever have, or that all
cognitive subjects must (ever) have conscious thoughts containing the concept BEING,
etc. Transcendental concepts, and so also the concept BEING (in the B-B sense) do not
have temporal priority; as thought-parts, they may not be acquired until later, or even
not at all.

For Aquinas, transcendentals are first in both ontological and cognitive order.
Cognitive perspective, however, is the basis and the starting point for Aquinas’s
reflections. Within priority in the cognitive order, we are going to distinguish
(following Tavuzzi) between the 'logical' and the 'psychological' order of priority. The
logical order of priority and posteriority

Q86. holds between different concepts on the basis of the relations of necessary inclusion and
exclusion that hold between the contents of these concepts: between the rationes that these
concepts express. Certain concepts contain as a necessary part of their content the contents of
other concepts which therefore need to be understood before they themselves can be

---

121 As Aertsen notices, Aquinas actually "develops his doctrine of the transcendentals from a cognitive
perspective" (Aertsen, 1996, p. 104). Epistemological reflections precede conclusions about the nature of
things in the world. For Aquinas this is exactly how we need to start our quest for knowledge: our only
contact with the world is through cognition, and so, a careful analysis of cognition is required as a first step
before we can start to investigate the world itself; "a reflexive analysis of human knowledge allows us to
175–83). (If Aquinas is really making conclusions as to what the world must be like on the basis of his
analysis of the cognizing subject, then this sounds quite similar to the approach taken by Kant. Obviously,
however, Aquinas still remains a realist. This is because, as we said above, cognition, for Aquinas, is to a
great extent passive. It has to start with the external senses which can only register stimuli from the external
world existing independently of any cognizers.)
understood: these other concepts which precede them in this manner are prior to them in the logical order. (Tavuzzi, 1987, p. 555)

The concept FOX is logically prior to the concept VIXEN, and, to use Aquinas's own example, the (transcendental) concept BEING is logically prior to the (transcendental) concept GOOD (Ens secundum rationem est prius quam bonum; DV 1, 1). When Aquinas says that ens est primum intelligibile secundum rationem (ST I 5, 2), he has in mind the logical priority of Being. He talks about logical priority of Being also in DV 1, 1, where he says:

Q87. that (…) to which [the intellect] reduces all its concepts, is being. Consequently, all the other conceptions of the intellect are had by additions to being. (DV 1, 1)122

The psychological order of priority and posteriority concerns the various necessary conditions which must be realized if an act of understanding or intellectual cognition is to be carried out at all (cf. Tavuzzi, 1987, p. 556). For instance, the capacity to apply concepts is psychologically prior to knowledge of theoretical physics. Being's psychological priority is what Aquinas refers to when he says that ens est primum quod cadit in apprehensione intellectu ("being is that which is first apprehended by the intellect," ST I 5, 2) without which "nothing can be apprehended by the intellect" (In I Sent 8, 1, 3). When Aquinas talks about transcendentals as being psychologically first, he does not refer to them as concepts. Rather, what he is talking about is the innate programming of

---

122 See also In I Sent 8, 1, 3; DV 21, 1; 21, 4, ad4; DEEE prologue; DP 9, 7, ad15; ST I 5, 2; 11, 2, ad4; ST I–II 55, 4, ad1; 94, 2 and more.
the human mind: transcendentals are psychologically first when they are understood as innate cognitive capacities or mechanisms.

Consider now how we could apply different kinds of priority in the case of sensation. For Aquinas, just as "being is the first intelligible" and "sound is the first audible" (ST I 5, 2), color is the first visible. The firstness of color does not consist in temporal priority. What a baby sees first is not color as such, not something that would be the content of the concept COLOR. Rather, what babies see first is their mother, or, say, a pink-ish moving object. Once we distinguish between an object which is the holder of a sensible quality, i.e., the substance that has a given accident, and the sensible quality itself, we notice that the temporally first object of sight is an accidental sensible (cf. above 1.1.2 and 2.3.2). 'Colored' on its own is not really a quality that is seen; we can only see particular colors, particular shades of pink, yellow, green, etc. There is no such thing as 'seeing something (merely) colored.' Nonetheless, the concept COLORED is logically prior ('first') to the concept PINK and to all concepts of particular colors. Anything that is pink is also colored. A thing can be pink only if it is colored, but it can be colored without being pink.

There is also another way to understand the priority of COLORED here. In what concerns vision, what really is first and the most fundamental in the cognitive order is not any concept understood as a thought-part. There can be cognition of colored objects without the subject possessing any concepts of colors. The most fundamental cognitive priority of COLORED consists in the fact that it is this exact feature of objects that is detectable by the sense of sight. A necessary condition of seeing a pink moving object is
the possession of the capacity—an innate mechanism—to see colors. In general, to figure out what is first in the psychological order is to know what is necessary for human beings to be able to engage in a given kind of cognitive activity and what constitutes the fundamental necessary condition for cognition. A similar reasoning can be applied to Aquinas’s claim to the effect that being is the first object of the intellect: it is also first in both logical and psychological order, and not in the temporal order.

(ii) Logical Priority of Being

The concept BEING is logically first in the sense that every other concept reduces to it. The concept BEING is logically prior to all other transcendental concepts, and to all universal concepts—all categories are created by adding to BEING.

Logical priority of the transcendental concepts is especially important in the context of the need for a starting point of science in face of the threat of the infinite regress. Aquinas's point of reference in his discussion of this matter is Aristotle. For Aristotle, science proceeds by means of arguments, or syllogisms. In order, however, for the conclusions of these syllogisms to have the value we want them to have, i.e., in order for them to be true, they must proceed from premises that are themselves "true, primary, immediate, better known than and prior to the conclusion" (PA I 1, 71b22). Some premises acquire their truth in virtue of being conclusions of prior syllogisms. In order
for the chain of reasoning not to go on forever, in order to stop the regress of
demonstrations, we need a set of premises which would be a starting point for science:
premises which are self-evident and don't need support of any kind. These are the first
premises of science. Aristotle explains:

Q88. Our own doctrine is that not all knowledge is demonstrative: on the contrary,
knowledge of the immediate premises is independent of demonstration. (The necessity of this
is obvious; for since we must know the prior premises from which the demonstration is
drawn, and since the regress must end in immediate truths, those truths must be
indemonstrable.) (PA I 3, 72b18-24)

Aquinas shares Aristotle's view on first principles as being the starting point in the
process of deliberation. He also follows Aristotle in another claim, that is, in the claim
that there is a special way in which we augment the scope of our knowledge: it is
through investigation into what things are by means of finding their definitions. The
stage of forming definitions, like that of deliberation, requires a reliable starting point.
As Aquinas says in his commentary on Boethius's De Trinitate,

Q89. In the speculative sciences we always proceed from something previously known, both
in demonstrating propositions and also in finding definitions (...). But it is impossible to go
on to infinity in this case, because then all science would cease, both as regards
demonstrations and as regards definitions, since the infinite cannot be traversed. So inquiry
in all the speculative sciences works back (reducitur) to something first given (prima). (InBoeth
6, 4)

The forming of definitions is the process in which we try to figure out what something
is. For Aquinas, to give a definition of something means to provide its genus proximum
and differentia specifica (cf. InBoeth 6, 4, etc), to say what genus it belongs to, and in what

123 Aquinas makes a similar point in DV I 1. Cf. also Aristotle, PA I 1; 71a12: "All teaching and all
intellectual learning come about from already existing knowledge."
it differs from other members of the genus. That is, the goal is to reduce the object
defined to a higher, that is, a more general concept.

What we need to figure out, therefore, is whether the need for a reliable starting
point in science and in the forming of definitions requires that we interpret first
principles and transcendental concepts as pieces of knowledge.

(iii) Digression: What are First Principles and Transcendental Concepts?

Aquinas's statement that in science "we always proceed from something previously
known" indicates something innate; in particular, it may be interpreted as suggesting
that there is innate propositional knowledge—first principles, and innate concepts that
are parts of the first principles. There are, indeed, other passages in Aquinas's work
which seem to support such an interpretation. In *De Veritate*, Aquinas says:

> Q90. We must give a similar explanation of the acquisition of knowledge. For certain seeds of
knowledge pre-exist in us, namely, the first concepts of understanding, which by the light of
the agent intellect are immediately known through the species abstracted from sensible
things. These are either complex, as axioms, or simple, as the notions of being, of the one, and
so on, which the understanding grasps immediately. In these general principles, however, all
the consequences are included as in certain seminal principles. When, therefore, the mind is
led from these general notions to actual knowledge of the particular things, which it knew
previously in general and, as it were, potentially, then one is said to acquire knowledge. (*DV
11, 1*)

First principles in the form of propositional knowledge, and transcendental concepts in
the sense of thought-parts are logically prior: everything reduces to them. They
constitute a reliable starting point for our scientific derivations and the forming of
definitions.
Consider, however, the following passage:

Q91. The agent intellect must exist prior to the *habitus* of first indemonstrable principles in order to be the cause of it. Indeed, the principles themselves are related to the agent intellect as certain of its instruments, because the intellect makes things actually intelligible by means of such principles (*QDA* 5)

The agent intellect, according to Aquinas, precedes in existence and is the cause of the knowledge of first principles. The intellect uses first principles as its instruments. This suggests that they are some kinds of mechanisms, principles of the functioning of our intellect. More interestingly and more surprisingly, as it turns out, first principles in some way depend on sensation:

Q92. It is clear that the only things we can know in the speculative sciences, either through demonstration or definition, are those that lie within the range of these naturally known principles. Now these principles are revealed to man by the light of the agent intellect, which is something natural to him; and this light makes things known to us only to the extent that it renders images actually intelligible; for in this consists the operation of the agent intellect, as the *De Anima* says. Now images are taken from the senses. So our knowledge of the above-mentioned principles begins in the senses and memory, as is evident from the Philosopher. Consequently, these principles do not carry us beyond that which we can know from the objects grasped by the senses. (*InBoeth* 6, 4)

Sensation is a necessary condition for intellectual cognition. Sensory experience activates certain patterns of functioning of the human intellect.

If we agree that it is useful to interpret them as mechanisms, we’d conclude that first principles as propositions are derivative; the propositions are only an expression of what first principles really stand for.
(iv) More on Psychological Priority

As we said, Being's psychological priority is what Aquinas refers to when he says that Being is the first intelligible (ST I 5, 2) without which "nothing can be apprehended by the intellect" (In I Sent 8, 1, 3).124

We have said that the psychological order of priority and posteriority concerns the various necessary conditions which must be realized if an act of understanding or intellectual cognition is to be carried out at all (the capacity to operate on numbers is psychologically prior with respect to the capacity to use calculus). The transcendentals, I'm going to claim, are psychologically prior in the sense that they correspond to innate mechanisms necessary for human cognizers to be able to think and to acquire concepts.

Let us go back to Aquinas's claim that being is the first intelligible. According to Wippel (Wippel, 2007, pp. 38–40), this idea indicates that on Aquinas's account some understanding of Being is available to every human subject endowed with an intellect. At the same time, Wippel admits, Aquinas certainly does not believe that everyone is a born metaphysician, so Being is certainly not temporally prior; it cannot be the case that we all have the concept of Being-as-Being, which is the subject matter of metaphysics. To solve this conundrum Wippel suggests that we distinguish two notions of being, being-as-being in its sophisticated sense as an object of metaphysics, and what he calls the ordinary pre-metaphysical grasp of reality, or of being, open to everyone.

124 The concept FOX is logically prior to the concept VIXEN, but it is not a psychologically first concept. When Aquinas talks about firstness of transcendental concepts in the sense of what we take to be psychological priority, he is not really referring to concepts but rather to how the mind functions, to certain capacities of the mind that are necessary for cognition.
This sounds plausible until we realize how exactly Wippel explains the so-called pre-metaphysical grasp of being (see Wippel, 2000, pp. 38–40). Wippel notices that for Aquinas the intellect is able to make any judgments at all, and so also judgments about the existence of things that it cognizes, only in its second operation (see 1.1.2 (v)), after it has produced concepts which represent essences of things. Once the intellect makes several such judgments, Wippel explains, "one will be in position to reflect upon this procedure and as a consequence to form in some vague and general fashion one's idea of reality or being meaning thereby that which is" (Wippel, 2000, p. 39). It is at this point that one will have arrived at what Wippel calls a primitive notion of being.

I do not think this is a good explanation of Aquinas's claims about Being as the first intelligible that I take to be crucial for Aquinas's account of cognition in general. It is true, Wippel does avoid the idea of ascribing to Aquinas the implausible view about temporal priority of Being. The view that Wippel ascribes to Aquinas, however, sounds rather trivial. Wippel himself admits that this notion of Being common to all thinking subjects is "still restricted to the kinds of being with which our knowledge begins, i.e., material and changing beings." But this shows that what Wippel has in mind is limited to Being as B-0, and so, it does not really concern Being in the B-B sense at all. It is hard to see how a claim about things in the B-0 sense, which constitute a small subset of possible objects of cognition, could be in any way crucial for Aquinas's account of cognition in general.

The problem with Wippel's interpretation seems to be that he takes Aquinas's claim that Being is the first intelligible as a claim about possession of non-transcendental
concepts, and just rejects the view that it is a claim about a metaphysical concept of being. This is probably why the only kind of priority of Being that Wippel considers is priority in the logical order. Again, he is right in suggesting that Being, for Aquinas, is logically first—this is exactly what Aquinas means when he says that "that to which [the intellect] reduces all its concepts is being" (DV 1, 1; cf. Aertsen, 1996, pp. 80–82). Aquinas means something else, however, when he says in addition that "without being, nothing can be apprehended by the intellect" (In I Sent 8, 1, 3) or that "the first thing conceived by the intellect is being; because everything is knowable only inasmuch as it is in actuality" (ST I 5, 2).

And so I suggest a different understanding of Aquinas's claim to the effect that Being is the first known, ens primum cognitum. For Aquinas, just as "being is the first intelligible" and "sound is the first audible" (ST I 5, 2), color is the first visible. This firstness of color does not consist in temporal priority. What baby Jane sees first is not color as such, not something that would be the content of the concept COLOR. Rather, what she sees first is her mother, or, because babies don't see their mothers as mothers, perhaps we should say that what baby Jane sees first is a pink-ish moving object. (If we apply Aquinas's distinction between three kinds of objects of sensation (cf. 1.1.2 (ii)), we notice that the temporally first object of sight is an incidental sensible.) 'Colored' is not a quality that can be seen on its own; there is no such thing as seeing something merely colored. Rather, we can only see particular colors, pink, yellow, green, etc. Nonetheless, the concept COLORED is logically prior ('first') to the concept PINK and to all concepts
of particular colors. Anything that is pink is also colored. A thing can be pink only if it is colored, but it can be colored without being pink.

There is also another way to understand the priority of color here. The most fundamental cognitive priority of color consists in the fact that it is this exact feature of objects that is detectable by the sense of sight. A necessary condition of seeing a pink moving object is the possession of the capacity—an innate mechanism—to see colors. In general, to figure out what is first in the psychological order is to know what is necessary for human beings to be able to engage in a given kind of cognitive activity and what constitutes the fundamental necessary condition for cognition.

The same interpretation can be made of Aquinas’s claim that being is the first intelligible: it is psychological priority that Aquinas has in mind. Consider again the following passages:

Q93. without being, nothing can be apprehended by the intellect (In I Sent 8, 1, 3); 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, and is primarily intelligible; as sound is that which is primarily audible. (ST I 5, 2)

As the passage from the Summa Theologiae suggests, the priority of Being consists in the fact that "everything is knowable only inasmuch as it is in actuality" (ST I 5, 2). 'To be in actuality' means 'to be actual,' and so, 'to be a Being.' Being turns out to be a necessary condition for knowledge or cognition in general. That there are colored things in the world makes vision possible. And this is because vision is a color-detecting mechanism; it consists in this innate capacity on the part of the subject to detect and react to colors.
The situation is similar with Being (i.e., what color is to vision, Being is to the intellect). The fact of there being beings, or the fact of the being-in-act, or of actuality of things in the world is what makes intellectual cognition possible. This is because the intellect is an innate Being-detecting mechanism. So to speak, Being is 'the quality' that the intellect is by its nature determined to detect and register. If the intellect cognizes any thing at all, then this 'thing' must be a Being. When we say that Being is the proper object of the intellect what we really mean by 'Being' is not what the intellect cognizes, but how it cognizes, what format it can read.

To say that Being is first in the psychological order, i.e., that it is a necessary condition for knowledge, is not equivalent to the trivial claim (similar to the claim that I ascribe to Wippel above) to the effect that there have to exist things in the world for us to be able to cognize anything at all. This claim would apply, surely, but merely to cognition at the level of the senses. Sensory cognition, for Aquinas, is passive. Its passivity is in fact a proof of metaphysical realism: we sense things in the world because these things exist (with the possibility of sensory illusions, but these, for Aquinas, are exceptions). This is not so with the intellect. First of all, the intellect is not passive (and it is plausible that it can actually create Beings). Secondly, as we have seen, transcendental 'existence' or Being is not limited to physical things existing in the external world. The scope of transcendental Being goes beyond 'things existing in the external world.' To be a being, in the end, is to be something intelligible. As Aertsen explains,

125 The claim about psychological priority of Being would become even more significant once we substitute Being with the other transcendents.
Q94. Being is that on the basis of which things are capable of being known by an intellect, it is the prerequisite condition for every intelligible object, for something is intelligible insofar as it has being. (Aertsen, 1996, p. 80)

And so, (today) we can say that we are genetically determined to detect and react to Beings and Unities, to individual substances and their individual properties. The intellect cannot operate at all unless it identifies something as possessing the transcendental properties. Still, we first acquire concepts of particular kinds of material beings, such as CAT, STONE, or MOTHER, that we cognize by means of our senses. Most people never think of things in the world as Beings in the B-B sense, and so, they never acquire the concept of transcendental BEING. (They may think of things in the world as individual existing things, and so acquire the concept BEING in the B-0 sense.) They have CAT-, STONE-, LOVE-thoughts, but not transcendental-BEING-thoughts. At the same time, transcendental BEING underlies all of these CAT-, STONE- and LOVE-thoughts: all these things can be cognized only because they exist in the B-B sense (and also, assuming that we would manage to explain other transcendentals as indicating additional features of the innate structure of the intellectual cognition, because they have an essence, are distinct from other things, are not divided, are intelligible, etc.). Anything that we cognize, and so, also material things in the external world, we cognize through Being and other transcendentals. We do it, though, without (at least at first) having transcendental concepts BEING or UNITY.

This goes against Pasnau's suggestion to the effect that for Aquinas human cognition begins with the understanding of the terms of first principles, that is, with the
understanding of transcendental concepts. First principles are grasped, according to Pasnau, "as soon as the nature of being is grasped" and the concept BEING is "what every intellect first conceives" (Pasnau, 2002, p. 326). According to Pasnau, Aquinas needs to posit this kind of immediate understanding of first principles and concepts that constitute them because otherwise human knowledge in general would have to start ex nihilo, and this is a view that Aquinas would certainly reject.

Notice that if Pasnau's interpretation works at all, it does so only because he at the same time ascribes to Aquinas the view that it is via divine illumination that we have this capacity to immediately recognize the truth of first principles. Without appealing to God, or to the supernatural, Pasnau claims, Aquinas would have "no way of explaining how we acquire our most basic concepts and recognize the truth of first principles" (Pasnau, 2002, p. 309). If, as Pasnau suggests, Aquinas endorses illuminationism, then he can supposedly claim that "the intellect begins life a blank slate, then quickly acquires the ideas [concepts] (being, one, good) that are the foundation for all other concepts" (Pasnau, 2002, p. 326). This, according to Pasnau, shows the limits of Aquinas's empiricism and naturalism.

On my interpretation, Aquinas's empiricism and naturalism are not threatened. Nor is there a need to suggest that Aquinas endorses illuminationism.

Consider the following passage:

Q95. 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, and is primarily intelligible; as sound is that which is primarily audible. (ST I 5, 2)
Being is psychologically prior in the sense that "everything is knowable only inasmuch as it is in actuality" (ST I 5, 2). 'To be in actuality' means 'to be actual,' and so, 'to be a being.' Being turns out to be a necessary condition for knowledge or cognition in general. That there are colored things in the world makes vision possible. And this is because vision is a color-detecting mechanism; it consists of an innate capacity on the part of the subject to detect and react to colors. The situation is similar with Being. The fact of there being beings in the B-B, or things existing in the transcendental sense of the term 'existence', things-in-actuality, in the world is what makes intellectual cognition possible. This is because intellectual cognition is an innate transcendental being-detecting mechanism. Being and other transcendentals are the innate programming of the intellect—they determine the way in which the intellect can cognize the world.

3.3 Collaboration of the agent intellect and the cogitative power

The issue we need to consider now is in what way Being and other transcendentals are necessary for cognition (and so psychologically prior), how, in what sense, they function as innate mechanisms, and in what way they enable human cognizers to think and to use concepts.
3.3.1 Being as the first intelligible means Being as a cognitive mechanism

Let us go back to the idea that, as Aquinas puts it in the *Summa Theologiae*, "being is the first intelligible as sound is the first audible" (ST I 5, 2: *ens (…) est primum intelligibile, sicut sonus est primum audibile*). We already agreed that 'the first intelligible' here should not be understood in the sense of temporal priority, that is, as the first concept that the intellect operates on. Also, what we want to focus on is not logical but psychological priority of Being. We said that what's psychologically prior in the case of vision is the innate capacity to react to colors; similarly, psychological priority of Being consists in the fact of our cognition having as its necessary and fundamental element an innate being-detecting mechanism.

Like color is the proper object of vision, being is the proper object of the intellect. So to speak, being is 'the quality' that the intellect is genetically determined to detect and register. If the intellect cognizes any thing at all, then this 'thing' must be a being. When we say that being is the proper object of the intellect what we really mean by 'being' is not what the intellect cognizes, but how (or by what means) it cognizes, what format it can read.

In another passage in the *Summa Theologiae*, Aquinas explains:

Q96. The first object of our intellect in our present existence is not being (…) of any sort, but being (…) viewed in material things, through which we come to a knowledge of all other
The way I interpret it is as follows. As a result of natural selection, we are now genetically determined to detect and react to individual substances and their individual properties (similarly as, for instance, our vision is genetically determined to react to sharp contours of objects and to detect edges; see below, esp. 5.2.4, where I suggest that an innate trait should be understood as a trait that is an evolutionary adaptation, inherited by the individual). This means that all of us, young kids included, perceive beings, unities, etc., if we perceive anything at all. We do it, though, without (at least at first) having the concepts BEING or UNITY. What we acquire first are concepts of particular material beings, such as CAT, STONE, or MOTHER, that we cognize by means of our senses. Most people never think of things in the world as 'beings' in the metaphysical sense. They have CAT-, STONE-, LOVE-thoughts, but not transcendental BEING-thoughts. But still, BEING in the B-B sense underlies all of these CAT-, STONE- and LOVE-thoughts: all these things can be cognized only because they are Beings in the B-B sense. We cognize all material things through transcendental Being (cf. above 3.2.1 and 3.2.2).

In general, instead of saying that human cognizers must innately possess concepts of BEING and UNITY because otherwise no cognition would be possible, we should say, as I already indicated, that they only 'react to things' as beings.

---

126 Cf. also ST I–II 94, 2; *InMetaph* I 2, 46; *InMetaph* IV 6, 605; *DV* 1, 1.
So far it may seem that I am trying to make a quite implausible claim to the effect that a person can believe that there is a material object in front of them without at the same time believing that there exists an individual object in front of them. Well, I certainly admit that there is something fundamental and first, and necessary for all cognition, and that this 'something fundamental' does concern the existence of things. For this reason, it is not surprising that one may be reluctant to agree that, say, a baby doesn't think that the cup in front of her exists. What I want to emphasize, however, is that this fundamental, first and necessary thing for cognition is not the possession of an innate concept BEING (nor the possession of any concept). The belief in existence is only secondary. And this is exactly my point: that BEING cannot be the first concept of the intellect. (And since it cannot be the first, it doesn't have to be innate.)

When Aquinas talks about the firstness of Being, he cannot mean that it is a necessary condition of cognition that we possess transcendental concepts as thought-parts. Most people never think of transcendental properties of things. And in those cases when they do, it doesn't seem plausible to claim that the possession of such concepts is crucial for the very possibility of cognition. Similarly, Aquinas cannot mean transcendental concepts as mental words: these are extremely difficult for human beings to acquire. Still, Aquinas does call 'being' the first object of the intellect. What he means by it, I want to claim, is that Being stands for a precondition for cognition, an innate mechanism; it is an indication of how our intellect is programmed to function.

How can we cognize something and not have the concept BEING? Well, first of all, we should notice that for Aquinas properly speaking there can be no such thing as a
mental word BEING. Aquinas follows Aristotle in his claim that (transcendental) being is not a genus (cf. *DV* 27, 1: *ens enim non est genus, sed multipliciter dicitur de diversis*; cf. also Aristotle, *PA* 2, 7: "what is is not a genus"); and *Metaph* II 3, 993b 23). As we have seen, for Aristotle, the genus-differentia definition, by means of which we discover what something is, requires that the object we consider be defined through its *genus* and a *differentia* (cf. above, the part on Logical Priority of Being, where I explain what this kind of definition involves). However, there is nothing that would lie outside of what is predicated of Being, because Being is predicated of everything that there is; it is somehow presupposed in all the ascriptions of properties (cf. Aristotle, *Topics* 4, 1, 121a19: "of everything that is, being and one are predicated"). Being has the broadest scope, it applies to everything, and so it cannot be derived from any higher concepts, because there are no such concepts. Therefore, transcendental Being is indefinable because there is nothing which can serve as a *differentia*; if it is indefinable, then it has no content and it cannot be a concept in the sense of a mental word. Pure transcendental existence is not something of which we could have a mental word. We think by means of concepts, and concepts point to essences of things, to *what kind of thing* we are thinking about. We are able to think about pure transcendental existence, but when we try to capture it in a definition, it turns out to be impossible: we immediately go beyond existence, and start talking about the essence (see above, 3.2.2).

Another way in which we could try to make more sense of the idea that there can be cognition of beings without the cognizing subject possessing the concept BEING is by considering a few examples. Imagine that one Monday morning John is in his office,
deep in his thoughts. Suddenly the door opens and Keira Knightley walks in. What is it that would strike John in such a situation? What concepts would his intellect use in response to this event? John would probably notice how beautiful Keira is, what she wears, he’d smell her perfumes, think how strange it is that she shows up in his office, etc. Would he, though, be struck by the fact that now there are not one, but two things possessing the property of transcendental existence, or two beings in the B-B sense in his office? (This certainly is a possible scenario, but, unless John is a philosopher studying the transcendental concept BEING, it seems quite unlikely that this would be a thought that occurred to him.)

Consider another scenario: the situation is similar as before, but it happens on a Tuesday. This time John notices a mouse wandering around his office. Again, John would notice that it’s a mouse, he’d notice its color and smell (although the latter maybe only in his imagination), and perhaps it would make him think about the cat he used to have in the past. But he would not think: oh, a being, an existing thing! And then, how likely is it that John would think about the two events, Keira entering his office on Monday, and the mouse wandering around on Tuesday, and compare them in his mind because of how similar they are? Yet in one very crucial respect they are very similar: on both days a being visited John's office. Most likely, however, when John perceives a mouse on the kitchen floor, he does not think of it as existing in the transcendental sense of existence. If he thinks of the mouse at all, he thinks of it in terms of concepts GRAY, or

127 Similarly, think of what happens when you look around your room. You do not really see beings, existing things. Rather, you notice books, chairs, cups, etc. Of course, you are able to reflect on that, and think how amazing it is that there is something (i.e., different beings) rather than nothing in your room. But in normal circumstances you simply do not do it.
MOUSE, or RODENT-EXTERMINATOR, but not in terms of transcendental BEING.

(This is exactly because it is not really possible to think of existence alone.)

This does not mean that some kind of knowledge of the mouse's existence is not there, somewhere in the background. It is, but (usually) not in a conscious and fully conceptual way, but rather only in an implicit way. (Being underlies all John's thoughts, but is not in itself what he thinks about.) In fact, it is not clear that this should even be called knowledge; perhaps it is some kind of knowledge by acquaintance, a kind of intuition of the thing as existing, but certainly not knowledge involving the concept BEING. It is not a conceptual kind of knowledge, but a precondition for any kind of knowledge-that to take place.

Maritain makes a similar point when he explains:

Q97. [Being] is the first of all concepts, because it springs in the mind at the first awakening of thought, at the first intelligible coming to grips with the experience of sense by transcending sense (…). [The] first idea formed by a child is not the idea of being; but the idea of being is implicit in the first idea which the child forms. (Maritain, 1948, p. 25; and footnote 12)

So, when we talk about a little child cognizing or perceiving an individual dog, it would be more correct to say—or at least this is the suggestion I want to make—that she reacts to the dog as to this individual thing. In a way she does not cognize it as an existing individual. She does not think of it as an individual existing thing, or a being. She thinks of it first as 'black,' 'moving,' 'barking'; then, listening to how others call the dog, she thinks of it as 'Yogi,' and finally, as 'dog.' The child reacts to a given thing because of the features it possesses, e.g. because it is red, or warm, or loud, etc., and in general because of the kind of thing it is. So we could say that the baby is cognitively sensitive to the
Aristotelian categories.\textsuperscript{128} (This is evidenced in the language-acquiring process: the first words a baby learns are words for different kinds of substances and their accidents.) And sensitivity to existence is a non-conceptual pre-condition of cognition.

Is it possible to provide any kind of evidence for my claim? We can conduct tests to see whether, how and when in their cognitive development babies respond to colors, faces, moving objects, or sounds, and we can conclude that to some of these things they react better (or sooner, etc.) than to others. And certainly, in our times scientists can come up with tests to verify almost anything one could possibly think of. But is it possible to test whether the baby reacts to beings?

One thing that scientific research does show is that from the very beginning of their lives infants are sensitive to contour, shape, size, patterns and colors of objects. They are sensitive to 'discriminable properties, objects, or events,' that is, to accidents and to substances that the accidents belong to.

Q98. From the very first months of life infants develop the ability to perceive and conceptualize in ways akin to that of adults. At the core of this process, and of cognitive development in general, is the ability to categorize or classify or, more specifically, the ability to group discriminable properties, objects, or events into classes by means of some principle or rule. (…) to categorize (…) is to respond to [things] in terms of their class membership rather than their uniqueness. (Rakison, Oakes, 2003, p. 3)

If infants can detect no patterns, that is, when the field of vision is homogenous, no cognition is taking place: "the eye movements of newborns [are] uncontrolled"

\textsuperscript{128} I don't want to deny that what underlies it all is that all these substances and accidents exist. Of course this is the case. But this is not what our cognitive faculties consciously focus on. I can only react to a bright green cat if it exists. But I react to it because of its color or because of its being a cat, and not because of its being a being.
(Rosenblith, Sims-Knight, 1985, p. 361), as if there was nothing there. For cognition to take place, our senses have to be able to hook themselves to properties of things. Even the most basic cognitive activity of infants consists of categorizing. And this belongs to the level of the whatness of the thing. We could say, therefore, that for human infants there is no direct or conscious cognition of being. At the same time, clearly, there is no cognition possible without it.

Consider one last example. Imagine a child who just burned her finger by touching a hot stove. The child's mother cries out 'Hot!' What's explicit in this situation is a sensory experience; what's implicit is that there is something out there, a being. The mother is

Q99. explicitly concerned with communicating the fact that the kitchen stove burns the finger. (...) [but] not explicitly concerned with communicating the fact that the kitchen stove is something-there [i.e., that it is a being], though knowledge of this fact is the least possible knowledge presupposed by and implied in knowledge of the fact that it burns the finger. Thus, our temporally first intellectual knowledge can be described as a knowledge whose explicit content is rooted in some sense experience or other, the implicit content of which is at least what can be expressed as something-there, i.e., being. (Bobik, 1965, pp. 4–5)

The child in this example reacts to the hot stove, which is an individual existing being. She reacts to it as to something that is hot, and not as to something that is a being. But of course, she could not react to it as hot if it didn't exist.

The idea that Being is a non-conceptual pre-condition of cognition is important for several reasons. It explains in what sense Aquinas can say that 'being is the first object of the intellect,' it points towards the transcendental as an innate operative
structure of the intellect, and also towards the necessity of collaboration between the intellect and the cogitative power.

3.3.2 From Essence as Thing to concepts as thought-parts

I suggested that Aquinas's claim to the effect that Being is the first object of the intellect indicates that the intellect is a transcendental being-detecting mechanism—it cannot cognize anything unless it is a being in the B-B sense (in this sense Being is the first intelligible), and to cognize beings it does not need any innate concepts. This answers problem P2, that is, the Being as the First Intelligible Problem. Aquinas's claim expresses not a view about concepts, but rather a view about how the mind works: it is a being-detecting mechanism. Transcendental Being, together with other transcendentals, are innate mechanisms, innate rules of functioning of the cogitative power.

The problem we still need to consider is P1, that is, the Concepts as Essences Problem, and also the two additional sub-problems (P5 and P6).

P1. Aquinas explicitly says that the proper object of the intellect is Essence. This at first may sound all right, given that, as we have learned, Essence can be treated as equivalent to the second transcendental, Thing, and as a transcendental it is interchangeable with Being. Also, Aquinas himself says, following Aristotle, that Being "is that which signifies the essence of a thing" (DEEE 1, 3). The problem appears, however, once we recall that on Aquinas's account essences (in the mind) are always universal in the sense that
numerically different things can have the same essence. Transcendental Being, on the other hand, is always individual (for Aquinas, everything that is intelligible is so in virtue of its own unique being-in-act, in virtue of its own act of transcendental existence). So, it seems that both cannot be the (first) object of the intellect. (In fact, because the intellect on its own has no access to individuals, saying that its proper object is transcendental Being does not sound very plausible anyway.)

The two additional problems are as follows:

P5. If the intellect can only operate on universal essences, then it seems that we would never be able to think about individual objects, about my friend John, or about my dog Yogi. There would be no such thing as intellectual cognition of an individual. (And indeed, there is for Aquinas no purely intellectual knowledge of individuals; knowledge of individual things is impossible without the collaboration between the intellect and the cogitative power.)

P6. We said that transcendentals are the most general concepts which do not really tell us anything about a thing that they apply to. So, from transcendentals to any kind of more interesting knowledge about things there is a long way to go. Even if we do have transcendentals as innate mechanisms of the intellect, this would not be enough to explain how we acquire, say, the concept DOG.

As we said above, the crucial fact for the purposes of this dissertation is that Essence and Thing can be used interchangeably. This fact is what helps us deal with problems P1 (Concepts as Essences), P3 (Transduction), and P5 mentioned above. We
should notice, first of all, that for Aquinas actually the five following terms, Essence, Quiddity, Thing, Nature and Form, are sometimes used interchangeably. Let us briefly describe each of the terms. In *DEEE* Aquinas explains:

Q100. the word 'nature' taken in this sense seems to mean the essence [*essentia*] of a thing in so far as it is related to the thing's own way of behaving, since nothing is without its own way of behaving. On the other hand, the word 'quiddity' is derived from what is referred to by a definition; whereas 'essence' means that through and in which a thing has its being [*esse*]. (*DEEE* 1, 4)

Q101. [Essence is] something common to all the natures through which different beings are placed in different genera and species (*DEEE* 1, 3).

According to the passage just quoted, Essence is that by which something is what it is, it is the whatness of a thing: the essence of a cat is that by which it is a cat. Recall, however, that this is how we defined Form, and not Essence. So, I'll keep this definition—that by which something is what it is, the constitutive ontological element of every existing thing—for Form ("that which makes a thing to be what it is (...) is also called 'form'," *DEEE* 1, 4). The most common understanding of the term Essence appears to be similar to that of Quiddity from the above passage from *DEEE* 1, 4: it is the nature of something as it is given in a definition (from now on, I take Essence and Quiddity to mean the same thing). On such an interpretation, Essence belongs to the language level of cognition. Essence in this sense is always a universal. It is not possible to give a definition of an individual *per se*. A definition of 'man' would, if it is correct, apply to John, but it is not possible to define John as such. Having flesh and bones is a part of the definition (and so of essence) of man. However, a definition of man does not make a reference to *this* flesh, or *these* bones. These belong to a particular man, to the nature of John, a particular
person, and not to the definition of man in general. Natures are always individual: they are always related "to the thing's one way of behaving" (DEEE 1, 4). While an Essence belongs to the level of language, a Nature is something metaphysical; it belongs to the ontological level.

'Nature' could be defined as an individualized essence as it appears in a thing. Things, like natures, are always individual. As Aquinas explains, Thing "expresses the quiddity or essence of the being" (DV 1, 1). This shows that 'Thing' is very close in meaning to 'Nature.' 'Nature' and 'thing' do not have the same meaning, however. A 'thing' is something that exists independently; and it is something which has a 'nature.' 'Thing' implies existence. 'Nature' is an indicator of essence. A Thing is always an individual object, a concrete here and now, but it is also an object which is an instance of a universal essence. (Also forms in things are always individual; only in the mind—as an object of cognition—can they become universal. In the world there are no universal forms.)

Let us go back to problem P1 (Concepts as Essences), i.e., to Aquinas's often repeated statement to the effect that the proper object of the intellect is Essence: the intellect cognizes only essences of things (as opposed to the senses which cognize only accidental aspects of reality).

If we take essence to mean the nature of something as it is given in a definition, then the intellectual cognition of a 'human being' would always require having a grasp of what really is the essence of humans. Indeed, on the common interpretation of Aquinas, when you intellectually cognize a person, this means that you must have a
mental grasp of *something* which is distinct from perceptible qualities and which seems very rich in content. This content is supposed to inform you what it really means to be a human being. It is supposed to represent the classical definition of a 'human being' in terms of its *genus proximum* and *differentia specifica*. In addition, it is commonly agreed that for Aquinas it is impossible to have this conceptual grasp of an essence of human beings without also having sensory experiences of individual people. So, human cognition, as Aquinas supposedly presents it, is very mysterious indeed: the senses provide information: pink, large, moving, making noises, etc. Then, somehow on *this basis*, the intellect thinks: 'homo sapiens.'

What I think could help us get rid at least of some of the mystery is to suggest that when Aquinas says that the proper object of the intellect is the essence, there are actually two cases that should be distinguished. These two cases correspond to the two ways in which Aquinas understands concepts: as *species intelligibilis* and as *verbum mentis* (cf. above, 3.1). What we need to consider now is the sense in which intelligible species and mental word can be said to stand for essences of things. Let us look at the following passage:

Q102. the intellect informed by the species of the object, by understanding produces in itself a kind of intention of the object understood, which intention reflects the nature of that object and is expressed in the definition thereof. This indeed is necessary: since the intellect understands indifferently a thing absent or present, and in this point agrees with the imagination: yet the intellect has this besides, that it understands a thing as separate from material conditions, without which it does not exist in reality; and this is impossible unless the intellect forms for itself the aforesaid intention. (*SCG* 1, 53)
The passage from the *Summa Contra Gentiles* suggests that while intelligible species is something which should be understood as the starting point of the intellectual operation, mental words are the end-result of the intellective cognition (cf. Panaccio, 2001, p. 188). Mental word, referred to as a kind of intention, is something produced by the intellect in the process of *understanding*. Mental words represent the nature of the object cognized as it is expressed in a definition. Aquinas explains that

Q103. whenever we understand, by the very fact of understanding there proceeds something within us, which is a conception of the object understood, a conception issuing from our intellectual power and proceeding from our knowledge of that object. This conception is (...) is called the word of the heart. (*ST I* 27, 1)

We have seen that it is plausible to interpret mental words as corresponding to concepts understood as conceptions (cf. 3.1.2). On the other hand, before the intellect produces this rich in content mental word, there is a prior stage of the intellect’s operation in which, as Aquinas says in the passage from *SCG* 1, 53, it is ‘informed by the species of the object.’ Clearly, since nothing material can affect the immaterial intellect, the species that informs it cannot be the sensible, but only the intelligible species. The intelligible species is a tool that the intellect uses for *thinking*. It stands for a concept understood as a thought-part. A concept just is a thought part; it is a mental sign that is supposed to symbolize the *whatness* of a thing of a certain kind. In order to be able to use a given concept just to think about something, there is no need for the more detailed knowledge of what the *whatness* of the thing really is.
Once we connect this new interpretation of intelligible species and mental words to Aquinas's understanding of Essence, we can infer the following. On the one hand, Essence can be taken to have its most common meaning, i.e., that of a definition. When in this context we talk about the intellect cognizing essences, we'd conclude that the end product of this operation is concept as a mental word, i.e., as a conception. But clearly, the conceptual grasp of an essence in the sense of a definition of something is not always the end product of our cognition; we don't always end up with an exhaustive understanding of what a thing really is. But, we could say that this is always in a certain way the goal toward which all cognition is directed.

On the other hand, we can take Essence in the sense closer to that of Thing. In this case, when it is said that the intellect cognizes essences, we mean it in the sense that the product of the intellect's cognition is a concept, or an intelligible species, as a thought-part. Here the concept CAT that you get when cognizing cats, even though it still points towards what constitutes the essence of cats, it does not require any grasp of what the essence of a cat is. One can entertain a cat-thought without knowing its essence or definition. Here the concept is only a basic tool for thinking about something. This solves the problem P2.

We still haven't solved the problem P5. For Aquinas, only individual things exist. And the intellect by itself does not have access to individuals (DV 3, 6). Because only individual things exist, the intellect seems to have no way at all to grasp the existence of something. In addition, we have said that the transcendent Thing is interchangeable with Being, that is, with something existing, with some individually existing thing. This
suggests that the intellect on its own could not operate on the transcendental:
transcendentals always indicate something individual. This is, as we shall see, where the
cogitative power enters the stage.

### 3.3.3 The need for the cogitative power

Aquinas admits that it is possible to consider essences of things independently of
whether those things exist. However, in usual cases of cognition, before we can cognize
*what* something is, we first affirm the thing's existence: "we cannot even ask what it
[something] is until we know that it exists" (ST I 2, 2). The detection of Being is the
necessary condition of cognition to take place. The most basic human cognitive function
is spontaneous perception of material world as something *hic et nunc*, as something truly
existing, real. We can see that this makes sense once we recall that although for Aquinas
objects of cognition (in the sense of cats, dogs, trees, and other things in the world that
we can cognize) are only potentially intelligible, still if it is possible to talk about *any*
objects of cognition at all, it is because those things are themselves *in act*. As Aquinas
explains,

> Q104. what the intellect comprehends is formed in the intellect, the intelligible object being,
as it were, the active principle, and the intellect the passive principle. (CT I 38)

---

129 My focus here is on cognition of the external world independent of human cognizers. The account that I
develop would have to be somewhat modified to apply to creations of our minds.
130 The passage continues: "That which is thus comprehended by the intellect, existing as it does within the
intellect, is conformed both to the moving intelligible object, of which it is a certain likeness, and to the
quasi-passive intellect, which confers on it intelligible existence. Hence what is comprehended by the
intellect is not unfittingly called the conception of the intellect." (CT I 38)
For something to be cognizable, it must be intelligible. And to be intelligible, for Aquinas, it means to be a Being in the B-B sense, to be in actuality, to (transcendently) exist. What first confronts the intellect in the psychological order is the being-in-act of the object cognized (cf. Tavuzzi, 1987, pp. 577–8). To cognize something is first of all to cognize it as a Being in the B-B sense, that is, to cognize it as existing (in the transcendental sense of existence). Existence, however, is not a property—it is neither a part of the essence of an object, nor one of its accidents. Our senses and the intellect cognize a thing's accidental and essential features, they cognize what a thing is; they do not really cognize that it is. And so, transcendental existence, or Being, is not something that could be detected by the external senses, and it is not something that the intellect could cognize on its own. As Krapiec puts it,

> Q105. Existence can be discerned, affirmed, ascertained, experienced, but this experience of existence cannot be communicated to anyone on the basis of the cognition of existence alone. Every communication of the intellectual experience of existence takes place by means of the conventional signs of natural language or some other means of cognitive communication, and all of this is connected with the conceptual apprehension of some sort of content. (Krapiec, 1991, p. 95)

There is no cognition without cognition of some content. And this means that any kind of actual cognition must go beyond Being, because Being on its own doesn't have any content. So, we have this basic and necessary cognitive function—the cognizing of existence—but it turns out that the external senses can't detect it, and the intellect cannot cognize it on its own. According to Aquinas, our human intellect on its own would acquire only "a universal and confused knowledge of things." The intellect on its own
has access only to essences of things. However, no real existing thing is constituted \textit{merely} by its essence. Existing things cannot be reduced to their essences. If a thing's existence \textit{were} reducible to its essence, then the cognitive process could be exhausted by the forming of universal concepts and by operations on those concepts. However, because the content of an existing thing is not exhausted by its essence, when we only formulate a universal concept about this thing, our cognition of it is never adequate. The intellect, therefore, \textit{always} needs to return to phantasms in order to acquire "a knowledge of what is proper" to each thing, in order to pass from "incomplete to complete knowledge" (\textit{ST} I 14, 6).

A phantasm, for Aquinas, is always "connected with the extended and time" because it is "a likeness of a singular thing which is 'here and now'" (\textit{InDMR} 2, 314; cf. also \textit{ST} I 14, 6). Things in the world exist always in the here and now, and this is accessible only to the cogitative power, a cognitive faculty which deals with the \textit{hic et nunc}, and which operates on phantasms. (The external senses also cognize things here and now, but they do it in a different way; they only cognize different accidents of things that exist here and now; what we need is a power which cognizes the bearers of these properties, because these are the primary objects of cognition for human being.)

The indispensability of the cogitative power is also visible once we reflect on transcendentals. As we said, from the ontological point of view, Existence—or Being—is a transcendental property of all things, and so (most likely) also of (both possible and) actual objects of cognition. When existence is taken from the perspective of the cognizing subject—it is a transcendental concept. The relevant fact about transcendental
concepts in general is, as we should recall (cf. above, in 3.2.2), that even though they are the most universal, in the sense that they apply to everything at all, go beyond the categories, and are limitless unlike 'regular' universal concepts, still, in each case the transcendentals apply uniquely to only one existing individual thing. The intellect on its own has no access to individual things. As it turns out, therefore, there would be no cognition at all of the transcendental properties of things if the only cognitive faculties we had were the intellect and the external senses. The reason why we can cognize such properties is that we also possess the internal sense of the vis cogitativa, which provides access to individual things.

3.3.4 Creation of concepts by the intellect together with the cogitative power

The cogitative power guarantees access to individual things, but it also brings with itself more than just the information about the thing's existence. Even though it is guided by the intellect, it is at the same time the highest among the sensory cognitive faculties. As a sensory power, the vis cogitativa operates on the information provided by the lower internal senses, and ultimately — by the external senses. The external senses connect human beings with individual things in the external world by registering their accidental features. In Aquinas's metaphysics, however, the existence of accidents is always secondary and dependent. Consider the following passage:
Q106. Matter is contracted by its form to a determinate species, as a substance, belonging to a certain species, is contracted by a supervening accident to a determinate mode of being; for instance, man by whiteness. (*ST* I 44, 2)

What this passage shows is basically that the structure of reality is such that cognition of accidents already indicates the way towards some universal concepts. Accidents are always of some individual substance (the nature of attributes is to inhere in substances), and the substance always belongs to a species—it is always *of a kind*. Consider an explanation of the cognitive process offered by Haldane:

Q107. we can come to know the nature or quiddity (the "what-it-is-ness") of a thing only by attending to and reflecting on its qualities or accidental forms. In this respect substantial form is not itself immediately observable as a sensible feature alongside color, shape or location. (...) What I see is a substance (Molly) with certain qualities some of which express its essential nature and others of which are extrinsic to this. However, whereas seeing the colour of her coat is a matter of sensory perception, seeing that the coat is of fur and that the fur is that of a cat involves understanding. In this respect judgements about the existence and nature of substances are always acts of intellection and not of mere sense-perception. (Haldane, 2003, p. 91)

This is all correct, except that Haldane does not mention the crucial intermediary stage in the process of cognition. On Aquinas's account, our sense organs are modified by the stimuli coming from the external world. Sensible species produced by the external senses are informational states representing features of things in the world. These states are *not*, however, that to which our concepts about the world apply, or from which our concepts derive their meanings (cf. Evans 1982, p. 227). We do not cognize the world as consisting of clouds of raw sensory data. Rather, we cognize the world as containing individual things. When, as Haldane says, 'what I see is a substance (Molly),' this means that the cogitative power is at work.
The cogitative power is guided by the agent intellect in the sense that it analyzes and transforms the output it receives from the external senses by means of the transcendentals. It cognizes individual things as things of certain kinds because it looks at the data provided by the external senses through the transcendentals of the intellect. This is what allows the cogitative power to see individual substances as things 'of a certain kind' in the data provided by the external senses.

We said that the first task of the cogitative power consists in its being affected by the existence of this here and now individual thing. The cogitative power's cognition of the world through the transcendentals is not limited, however, to the first transcendental, that is, to Being (and so, to the affirmation of the existence of an individual thing). Rather, the cogitative power experiences things in the world through all the transcendentals. As it was discussed in 3.2.2, for Aquinas, anything that exists is a being and anything that is a being, has an essence, it is of a kind. One implies the other, being implies essence. Apprehending something as existing already implies that the intellect in collaboration with the cogitative power will look for the thing's essence and formulate concepts. The essence does not have to be immediately known, but that the thing is of a kind is already implied by the fact that it is experienced as existing. Existence, or Being, as the first transcendental immediately indicates/imply all other transcendentals.

What's most important for our purposes here is that Existence implies Essence, that is, being of some kind, and also being an individual thing, a unity (One) that is separate from all other things (Something). (In order to simplify things we limit our discussion at this point to the first four transcendentals.)
Transcendentals are in a sense innate rules that determine how the cogitative power proceeds. Each transcendental indicates something about the function of the cogitative power. When the cogitative power receives a cloud of stimuli registered by the external senses, it will put them together according to a certain pattern in such a way that it ends up cognizing an individual object, a being, this particular thing here and now, having this specific essence—and so, a being that is of a kind; something that is separate from other things, that has no contradictions within itself, that is, that has no parts that would not be that thing, etc. (so we have here esse, essentia, unum, and aliquid). The rules according to which the cogitative power works are responsible for the fact that we will cognize an individual cat, and not the cat's tail together with a piece of carpet and the tip of a knife. The rules are determined, like other the instinctive or innate mechanisms, by nature—or, as we would say today, by natural selection. We are just programmed to cognize cats and not their random parts, similarly as our vision is determined to make us see objects distributed in three-dimensional space, and similarly as the sheep is determined to see an enemy in a wolf.

The end-product of the working of the cogitative power, of its cognition of an individual thing as an individual substance of this kind, is a phantasm of the highest kind. Once this representation of the individual substance is produced, the intellect, then, can create a concept—a thought-part, a mental symbol of that substance (cf. 3.1). That concept is a tool by means of which the subject can relate to the individual-distinct from others-existing thing-of a certain kind that it cognizes.
Because the intelligible species is a symbol that represents the individual substance, it can also function as a universal concept of the intellect (it points to the universal essence expressing the kind that the given object belongs to). The concept as a thought-part stands for the transcendental 'Thing' which, as we said, is an instantiation of a universal Essence (cf. 3.3.2). At this stage we don't yet know what this essence is—for this a more sophisticated process of understanding is needed. But what we do have here is an intelligible species by means of which we can think about the object.

The intellect is a purely conceptual faculty, and the cogitative power is the faculty of recognition that applies concepts to perceived individuals. (When, for instance, a person sees an object as green or recognizes that it is green, that's what the cogitative power does.) When the cogitative power cognizes an individual substance, if it is a substance of a kind previously known, the intellect (together with memory) allows the cogitative power to classify the object cognized as a thing of that previously known kind. The cogitative power is an innate mechanism which allows the subject to recognize the thing of a certain kind based on a certain kind of experience, assuming that the subject already got a certain concept in question. If the cogitative power encounters an object of an unknown kind, it is ready for the intellect to produce a new symbol—a new concept that will represent it. These symbols are (most often) accompanied by words of a natural language.
(i) Where do concepts come from?

The cogitative power can be interpreted as an innate and instinctive mechanism, which is necessary for the cognizing subject to be able to cognize something as existing. This affirmation of a thing’s existence is, as we said, a precondition for any kind of cognition to take place. It is the first, pre-intellectual stage of cognition.\textsuperscript{131} (This is also what Tavuzzi calls a preliminary grasp of being; cf. Tavuzzi, 1987, pp. 577–8.) This grasp of the thing’s actuality is pre-conceptual in the sense that during this first stage of the cognitive process there is no formation of any sign of the thing, no image, and no concept at first.

Aquinas says:

\begin{quote}
Q108. That which is thus comprehended by the intellect, existing as it does within the intellect, is conformed both to the moving intelligible object, of which it is a certain likeness, and to the quasi-passive intellect, which confers on it intelligible existence. Hence what is comprehended by the intellect is not unfittingly called the conception of the intellect." (CT I 38)
\end{quote}

The intellect plays a creative role in concept acquisition, but it is still far from clear where it gets its concepts. We agreed that abstraction does not work (at least abstraction on the purely empiricist understanding of the term); the intellect does not abstract any universal essences of things from the phantasms. Neither, however, are intelligible species innate. One way in which we could imagine Aquinas answer the question of the origin of concepts and in general to substantiate his account of cognition is by taking a

\begin{footnote}
The intellect’s first operation, i.e., the ‘understanding of indivisibles,’ is already directed at the thing’s nature (cf. InDT 5, 3).
\end{footnote}
cure from Nicholas of Cusa (1401-1464), who was both thoroughly familiar with the Aristotelian-Thomistic philosophy, but also heavily influenced by Plato and Neoplatonism.

According to Cusanus, in order to explain the functioning of the human mind, it is necessary to understand its relationship to the mind of God. God's cognition consists in creating things from nothingness. By thinking, God creates things according to divine ideas, or exemplars. The human cognition, on the other hand, is described as 'vis assimilativa' (cf. Nicholas of Cusa, 1979, VII). It is also an active power of cognition, but this power is limited; it is not creative in the same way in which God's cognition is creative. Human beings, by cognizing, only create similitudes of things. By thinking, the human cognizer assimilates the ideas in her mind to (forms of) things in the external world. Still, the divine and the human minds are closely related because the human mind is the image of God's mind. But also, the human ideas or concepts are closely related to God's ideas (exemplars) in virtue of the fact that human ideas are produced as likenesses of things, and things are creations of God's exemplars.

We find indications of a similar view in Aquinas's own texts:

Q109. Whatever derives its act of existence from God has an idea in God. (DV 3, 5, otc1)

Q110. Every essence is derived from the divine essence. Therefore, whatever has an essence has an exemplar in God. (DV 3, 5, otc2)

Q111. The likeness of things in the divine intellect is one which causes things; for, whether a thing has a vigorous or a feeble share in the act of being, it has this from God alone; and because each thing participates in an act of existence given by God, the likeness of each is found in Him. (DV 2, 5)

---

132 I am grateful to prof. Jay Reuscher for bringing this suggestion to my notice.
According to Aquinas, anything that exists was created and is continuously sustained in being by God. Existence is equivalent to intelligibility. Anything that is intelligible, as the passages from DV 3, 5 and 2, 5 suggest, has an exemplar in God's mind. God's mind contains forms of everything that could ever be cognized.

The human mind, which is created as the image of God's mind, could be said to contain potentially the forms of every object that God could or would create and that He would plan for human beings to be able to understand or to think about.

We said that for Aquinas the intellect is immaterial, which means that it is pure potentiality. The immateriality of the human mind can be said to be a potentiality for creating similitudes of the exemplars. It is a potentiality to create a mental symbol for any possible object of thought offered to it by the cogitative power. Mental symbols, i.e., intelligible species produced by the intellect, and at the same time similitudes of divine exemplars, are assigned to individual substances cognized by the cogitative power. When the cogitative power cognizes an individual substance of a kind, this activates a pattern in the intellect: a symbol is assigned that allows the subject to think about the substance of this kind.

Even though all this may sound like some kind of neo-Platonic nativism with concepts of things pre-existing in God's mind, this kind of interpretation is not necessary. On the contrary, this way of speaking may still be an expression of Aquinas's naturalism and his metaphysical realism. Aquinas's view is that the objects that human beings cognize are not creations of the human mind. The human cognition is dependent on both
how our cognitive faculties are structured and on what the external world is like.

Aquinas’s conviction to the effect that the world was created by God (and that it is continuously supported in existence by its Creator) can be translated into the claim that the external world exists independently of the human cognizers.

This kind of solution would in fact fit well within the contemporary brain theory according to which cognition can be explained in terms of specific experiences activating specific neural patterns in the brain. Perception of a given object triggers certain cells, in a specific location in the brain, to start firing together and to form a new neural pattern (cf. below, part 5.3.5). Activation of a given neural pattern constitutes thinking a thought containing a specific thought-part, and so, it constitutes the occurrence of a given concept. Whenever a given neuronal structure fires up, the organism entertains a given representation.
CHAPTER IV
FODOR: A CONTEMPORARY NATIVIST

4.1 Fodor's Computational Theory of the Mind

Jerry Fodor supports the Computational Theory of Mind, or CTM (Fodor 1975, 1981b, 1998a). The main theses of the CTM can be summarized as follows:

T1: intentional mental states are relations to mental representations which belong to a language of thought (RTM + LOT)
T2: thinking is computation
T3: meaning is information (Fodor, 1998a, p. 12)

4.1.1 RTM + LOT

CTM is a version of the Representational Theory of Mind (RTM) according to which intentional mental states, such as beliefs and desires, should be understood as relations to mental representations referring to properties and states of affairs in the outside world (cf. Fodor, 1981b, p. 259).
On RTM, to believe that it is raining is to be in a particular relation (characteristic of the attitude of belief) to a mental representation whose semantic value is "it is raining"; to hope that it is raining is to be in a different relation (characteristic of the attitude of hoping rather than of believing) to a mental representation with the same semantic value. Intentional processes, e.g., thinking, reasoning or deciding, are processes involving the manipulation of mental representations.

RTM, Fodor believes, is a very useful tool of predicting and explaining intentional behavior of an agent (e.g., how the agent makes decisions, etc.), because it allows us to assume that the agent has means for representing to herself her own behavior as being of a certain kind. Fodor explains:

Q112. If you want to know how the organism will respond to an environmental event, you must first find out what properties it takes the event to have. (...) If you want to know why the organism behaved the way it did, you must first find out what description it intended its behavior to satisfy; what it took itself to be doing. (Fodor, 1975, p. 55; see also pp. 30-1)

The prediction of behavior depends on "the salience of the proximal stimulus in the causation of behavior" (Fodor, 1975, p. 55), that is, on what the agent takes the stimulus to be, or how the agent represents the stimulus. The 'proximal stimulus' is really "a proximal representation of the distal stimulus" (Fodor, 1975, p. 55). The explanation of behavior is a function of the 'proximal response,' that is, of how (under what description) the agent represents to herself her own action.

---

133 It is also a useful tool for explaining what Fodor calls 'concept learning' and the phenomenon of perceptual integration (Fodor, 1975, p. 56).
For Fodor, the representational system that turns out to be so useful for making decisions (and for explaining intentional behavior), "presupposes a language"; "it must share a number of characteristic features of real [natural] languages" (Fodor, 1975, p. 31). This is because

Q113. representation presupposes a medium of representation, and (...) there is no internal representation without an internal language. (Fodor, 1975, p. 55)

### 4.1.2 Language of Thought

Because the internal representational system is language-like, also the mental representations involved in the tokening of intentional states are best understood, according to Fodor, as similar to linguistic symbols, that is, to the words and sentences of a natural language such as English. Sentence-tokens are concrete strings of physical marks with a definite spatio-temporal location and distinctive causal powers, and also with abstract propositional contents reflected by their grammatical structure. Beliefs are sentence-like in the sense that they also exhibit something like syntactical structure. And they also can be assigned abstract propositional content that is somehow reflected by their formal syntactical structure. Beliefs are physically realized in the brain; they can be understood as brain-code tokens, as some kind of patterns of neuronal activity (and perhaps as traces of such patterns for those beliefs that aren't presently occurring).
On Fodor's proposal, mental representations belong to a *non-natural* innate language known as the language of thought, or LOT.\textsuperscript{134} (One of the reasons why LOT is a non-natural language is that LOT is a prerequisite for learning *any* natural language; if so, LOT itself cannot be a natural language.\textsuperscript{135} Also, it seems that a natural language could not be a medium of thought. If we agree that non-verbal animals, that is, animals which do not have a natural language, are capable of thought, and that thought requires computation, then we have to ascribe to animals an internal representational system which, by definition, cannot be a natural language; cf. Fodor, 1975, p. 56.) And so, believing that it is raining involves tokening in one's mind a sentence that is the analogue of the English sentence 'it is raining,' but that actually is a sentence in a non-natural mental language, LOT.

\textbf{i. LOT vs. Natural Language}

A natural language, such as English, consists of a finite set of words constituting its vocabulary, and a finite set of grammatical or syntactic rules that are employed in combining words to make complex structures such as phrases and sentences. Because the syntactic rules are recursive, i.e., they can be applied over and over again, the users of the natural language have the possibility to come up with an infinite set of distinct sentences. Language of thought, according to Fodor, is productive and creative in a similar way. There is a finite set of primitive concepts (Fodor, 1998a, p. 27; cf. Fodor, \textsuperscript{134} Fodor's LOT is innate in the sense that it is not a learned language, cf. Fodor, 1975, p. 67 and p. 97.\textsuperscript{135} Cf. Fodor, 1975, p. 79: "Learning a (first) language involves formulating and confirming hypotheses about the semantic properties of its predicates. (...) Organisms capable of learning a language must have prior access to some representational system in which such properties can be expressed."
1981b, p. 262: "There is a (finite) basis of lexical concepts"), i.e., of symbols of LOT, and a finite set of recursive syntactic rules (Fodor, 1998a, p. 27) for combining these symbols to form an infinite number of more complex structures of sentences of LOT.\(^{136}\)

Words and sentences of natural languages have meaning. The meaning of a sentence is exhaustively determined by its syntactic structure and the meaning of its component words\(^{137}\). Similarly, the language of thought also has a combinatorial syntax and semantics (cf. Fodor, 1975, p. 32).\(^{138}\)

LOT, according to Fodor, is "a representational system of very considerable richness" (Fodor, 1975, p. 31). Its rich expressive power allows it to express any meaning that English is capable of expressing: "nothing can be expressed in a natural language that can't be expressed in the language of thought" (Fodor, 1975, p. 84). This follows from the fact that understanding a sentence of English involves tokening a symbol of LOT that has the appropriate meaning\(^{139}\). For every sentence of the natural language that we are capable of understanding, there must be a sentence of LOT that has just the

\(^{136}\) Fodor does not provide support for the claim that the set of primitive concepts is finite. It seems, however, that he'd agree with the account he describes as empiricist according to which the set of concepts potentially available to an organism is the closure of the primitive concepts under the combinatorial mechanisms (Fodor, 1981b, p. 277). The primitive concepts, on this account, "are the possible outputs of the sensorium across the range of inputs that it responds to" (Fodor, 1981b, p. 264). For this reason, if we could specify all the possible outputs of the sensorium, we would know what primitive concepts the organism can potentially entertain. Both the set of potentially available primitive concepts and the combinatorial mechanisms are innately fixed by the biological endowment of the organism. And so, Fodor concludes, "the space of concepts potentially available to any given organism is completely determined by the innate endowment of that organism" (Fodor, 1981b, p. 277).

\(^{137}\) This does not hold for idioms. Idioms are often exceptions; even though they are phrasal, they don't have a logical syntax (Fodor, 1981b, p. 264). The meaning of non-idiomatic expressions is also influenced by other things such as context, intonation, hermeneutic abilities of the interpreter, non-linguistic behavior of the addressee, etc.

\(^{138}\) For the discussion of how similar LOT is to natural languages see chapters 2 and 3 of Fodor's 1975.

\(^{139}\) Cf. Fodor, 1998a, p. 2 where Fodor suggests that "word meanings just are concepts."
same meaning.\textsuperscript{140} In fact, LOT is "capable of expressing any concept that the organism can learn or entertain" (Fodor, 1975, p. 97; so really, LOT seems to have a richer expressive power than any natural language).

4.1.3 Computation

Fodor's CTM is a combination of RTM with the idea that "much of 'higher cognitive behavior' is rule governed" (Fodor, 1975, p. vii), or, in other words, that "thinking is computation\textsuperscript{141}" (Fodor, 1998a, p. 9). So, more precisely, on CTM, intentional states are computational relations to mental representations, and intentional processes involve manipulation of mental representations by means of computation. A computational operation is "an operation defined for (internal) formulae" (Fodor, 1975, p. 76). Token mental states are "relations between the organism and formulae (...) in the internal code" (Fodor, 1975, p. 75), that is, between the organism and simple or complex symbols of LOT representing properties and states of affairs in the outside world. Symbols of LOT have both semantic and syntactic properties. They are syntactically structured physical objects with semantic properties (Fodor, 1998a, p. 10).

\textsuperscript{140} More precisely, what Fodor seems to mean here is that both the LOT symbol for 'cat' and the word 'cat' have the same meaning because both of them are symbols and both carry the same information. At the same time, Fodor also talks about the LOT symbol for 'x' as actually being the meaning of the word 'x.'

\textsuperscript{141} Fodor believes that currently the best models for explaining behavior, concept learning and perceiving are those models which treat these phenomena as computational (cf. Fodor, 1975, p. 51). If this is really the case, this could serve as a support for the claim that RTM is the theory of mind that we should indeed choose, because computation, as Fodor very often emphasizes, presupposes a medium of computation—a representational system (cf. e.g., Fodor, 1975, p. 27); computation requires RTM.
The mind, on CTM, is a symbol-manipulating device (Fodor, 1983, p. 39). Processes of reasoning consist of the mind taking syntactically structured symbols as input and producing syntactically structured symbols as output. Because the relation between mental representations (computational elements) is taken to be syntactic, the mind's functioning can be explained "by analogy to the organization of idealized computing machines" (Fodor, 1983, p. 38); the mind can be considered a Turing machine (cf. Fodor, 2000, p. 13). Fodor explains:

Q114. It is feasible to think of (...) a system as a computer just insofar as it is possible to devise some mapping which pairs physical states of the device with formulae in a computing language in such fashion as to preserve desired semantic relations among the formulae. (Fodor, 1975, p. 73)

Like a computer, the mind can write symbols in its memory, compare them, retrieve them from memory, delete them, etc.

The important feature of computation, Fodor emphasizes, is that it reliably preserves semantic values of the symbols (Fodor, 1998a, p. 11). Similarly, the processes executed by the mind are semantically or logically coherent in that the output usually 'makes sense' given its input and the information stored in its memory. Fodor acknowledges at the same time that there are important differences between minds and Turing machines. The latter are "very simple devices" (Fodor, 1983, p.39), not "responsive to the flow of environmental events"; they are "closed computational

142 Any syntactic relation can be computed by some Turing machine; cf. Fodor, 2000, p. 30.

143 Fodor explains that claims to the effect that a certain concept "applies to a certain thing, are always susceptible of evaluation in such semantical terms as satisfied/unsatisfied, true/false, correct/incorrect" (Fodor, 1998a, p. 24). Computational manipulation of symbols preserves these semantic values.
systems" (Fodor, 1983, p. 39). Minds, on the other hand, cannot be closed systems, because their main function is to provide the subject with the information about the world. On the contrary, there must be a part of the mind, or a sub-system, that registers the information incoming from the external environment, and translates it into "mental symbols in whatever format cognitive processes demand of the representations that they apply to" (1983, p. 39).

4.1.4 Meaning is information

Information is a common phenomenon that we find in nature, wherever causes leave effects in a lawful manner. We can talk about information where that is "a correlation between two things that is produced by a lawful process (as opposed to coming about by sheer chance)" (Pinker, 1997, p. 65). For instance, we can say that the rings on a tree stump carry information about the age of the tree because the number of the rings lawfully correlates with the tree's age. So, in more general terms, we can talk of information whenever there are some kinds of patterns that can potentially be perceived as representations. In the case of the tree, the rings on the stump are the pattern, and they represent, or carry information about, the age of the tree. Information is that which the pattern represents.
On CTM, the meaning (or content) of mental representations or of the symbols of LOT consists in the information that they carry\textsuperscript{144}. Mental representations obtain their identity, in part, by carrying information about aspects of the environment. As Fodor explains, "what bestows content on mental representations is something about their causal-cum-nomological relation to the things that fall under them" (Fodor, 1998a, p. 12). The idea here is that intentional contents of mental states (that is, what they represent) are constituted by the nomic (i.e., lawful) relations these items bear to elements in the surrounding world (having a concept is "constituted by being in some sort of nomic, mind-world relation," Fodor, 1998a, p. 146). So, for instance, "what bestows upon a mental representation the content \textit{dog} is something about its tokenings being caused by dogs" (Fodor, 1998a, p. 12).\textsuperscript{145}

4.2 Fodor's account of concepts

Concepts, for Fodor, are mental entities, species of representations. They are constituents of cognitive mental states (Fodor, 1998a, p. 6), parts of propositional attitudes, such as

\textsuperscript{144} What is special about information with respect to language and thinking is that information is \textit{manipulated} or \textit{processed} by means of computation which involves preserving semantic properties of symbols. In addition, because symbols of LOT are physically embodied, they can function as mental causes and effects (cf. Pinker, 1997, p. 66).

\textsuperscript{145} As it is commonly known, the causal theory of content faces several serious problems. For instance, it may happen that a representation (a concept) is caused by something it does not represent: my concept \textit{DOG} may occur when I see a dog picture, or a well-disguised cat (cf. Prinz, 2004, p. 91). An advocate of the causal theory of content has to somehow explain that my concept \textit{DOG} refers only to a certain subset of the things that reliably cause it to be tokened.
beliefs, desires, thoughts, and the like (Fodor, 2000, pp. 14–15). A concept, Fodor says, is such a species of mental representation "which expresses a property" (Fodor, 1981b, p. 259).

Fodor distinguishes two kinds of concepts: lexical and phrasal. He explains that lexical concepts, such as GREEN or BACHELOR, are "expressible, in English, by an open sentence with a morphologically simple—viz. a monomorphemic—predicate term." The concept GREEN is expressible by the open sentence '… is green.' The concept DOG expresses the property 'dogness' in the sense that it applies to dogs (Fodor, 1981b, p. 259) and that it is the mental representation "normally evoked by utterances of the form of words '… is a dog'" (Fodor, 1981b, p. 260). A phrasal (or complex) concept, such as, e.g., the concept LIVES IN CHICAGO AND EATS MANGLEWORTS, or IS UNCOMMUNICATIVE, is "expressible (...) by an open sentence with a morphologically complex predicate term" (Fodor, 1981b, p. 261).

The set of lexical concepts of any natural language is finite (cf. Fodor, 1981b, p. 261). On the other hand, there are infinitely many phrasal/complex concepts, acquired by the application of the recursive constructive procedures to lexical concepts (cf. e.g., Fodor, 1981b, p. 262). Lexical concepts are constituents of phrasal concepts. As Fodor puts it, "there are infinitely many phrasal concepts which are acquired by the application of constructive procedures to a basis of lexical concepts" (Fodor, 1981b, p. 261). Phrasal concepts, therefore, should be reducible to lexical concepts. The semantic properties of phrasal concepts depend on the semantic properties of lexical concepts. "A sufficient condition for having the concept C," Fodor explains, "is being able to think about
something as (a) C (being able to bring the property C before the mind as such)" (Fodor, 2008, p. 138). 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. Having a concept, Fodor explains, is constituted by being in some sort of nomic, mind-world relation (Fodor, 1998a, p. 146). A concept C carries information about a property P if C is under the nomological control of P. To say that C is under the nomological control of P means that instances of P reliably cause Cs to be tokened, or that occurrences of a concept C reliably correlate with a property P. The latter is the case if and only if the concept C shows up, as a matter of natural law, only when instances of P are present in one's surroundings.

In his *Concepts* (Fodor, 1998a, pp. 23–39), Fodor lists five features that summarize what he believes to be non-negotiable conditions that concepts must satisfy on any plausible theory of concepts. These non-negotiable conditions are as follows:

i. Concepts are mental particulars.

Concepts are parts of thoughts, and, for Fodor, as we said above, thoughts themselves are mental particulars (cf. Fodor, 1998a, p. 23). Both thoughts and concepts are the kind of things that can function as mental causes and effects. (Fodor is opposed to the view according to which concepts are capacities. Also, he wouldn't hold that thoughts

---

146 Fodor’s definition of concept possession in terms of being in some sort of ‘nomic, mind-world relation’ certainly does not cover all the cases of having a concept. For instance, my having the concept SANTA CLAUS cannot be explained by suggesting that it was caused by Santa. Fodor would not deny that we possess many concepts through mostly linguistic interactions with other people (e.g., through what is called ‘reference borrowing’; cf. Sterelny, 1989, p. 129). He would be satisfied, it seems, if his account worked at least in some cases and applied to the possession of such concepts as DOG, RED or DOORKNOB.
or concepts, as they occur in the minds of thinkers, are abstract entities—abstract entities, such as numbers or propositions, cannot function as causes of behavior; and concepts, according to Fodor, "satisfy whatever ontological conditions have to be met by things that function as mental causes and effects" (Fodor, 1998a, p. 23).)

Even though the claim that concepts are mental particulars is useful because it helps to explain their causal powers, at the same time, it is also problematic. If my concept DOG is a particular mental item in my head, then how could I have different thoughts about dogs? What helps is the fact that for Fodor concepts, because of their function, that is, concepts as symbols, satisfy a type/token relation. As pieces of a brain-code, they are particulars: physical objects that can have causal powers. On the other hand, concepts as symbols, that is, in virtue of the fact that they represent, belong to abstract types. Because of this feature they are repeatable, can be shared, and can function as categories.

ii. Concepts express categories.

Concepts represent; they stand for things, they apply to things in the world (cf. Fodor, 1998a, p. 24). Things in the world 'fall under' concepts. Concept A is different from concept B if they differ in their intentional content, i.e., if they refer to different things. As we know from Frege, concepts cannot be individuated by intentional content alone. In his Concepts, Fodor suggests that the second parameter of concept individuation is the concept's syntactic structure, that is, its form: "a mental representation is individuated by its form and content" (Fodor, 1998a, p. 27; cf. also Fodor, 1998a, pp. 37–39).
iii. Compositionality

Concepts are the constituents of thoughts, and also (often) of other complex concepts. Complex mental representations inherit their contents from the contents of their constituents (Fodor, 1998a, p. 25). We can entertain an infinite number of distinct thoughts because we combine concepts compositionally (cf. Prinz, 2004, pp. 12–3). As constituents of thoughts, concepts help explain both productivity and systematicity of thoughts. Thinking is productive in the sense that a person can produce and understand infinitely many distinct beliefs (infinitely many complex expressions). Systematicity of thought can be seen in the fact that the ability to entertain one thought implies the ability to entertain another (it consists in the ability to combine the same expressions with the same meanings in different syntactic compounds). As Fodor explains,

Q115. There are infinitely many beliefs because there are infinitely many thoughts to express their objects. There are infinitely many thoughts because, though each mental representation is constructed by the application of a finite number of operations to a finite basis of primitive concepts, there is no upper bound to how many times such operations may apply in the course of a construction. (Fodor, 1998a, p. 27)

iv. Quite a lot of concepts must turn out to be learned.

In addition to the lexical/phrasal distinction mentioned above, concepts are also divided into primitive and complex. A concept is complex if it has an internal structure, i.e., if it has semantically interpretable components. A primitive concept is simple, unstructured: it does not have any semantically interpretable components. The
set of primitive concepts is smaller than the set of lexical concepts (cf. Fodor, 1981b, p. 263). Complex concepts can be either phrasal or lexical. A concept is both lexical and complex if it is expressed by morphemically simple predicate term (if it is lexicalized, that is, if it can be found in a dictionary), and still has an internal structure, that is, it could be reduced to primitive concepts. On a typical empiricist view, the concept DOG would count as lexical and complex: it is, supposedly, reducible to a set of primitive concepts such as FURRY, BARKING, GRAY, SMALL, etc.

Primitive concepts, out of which complex concepts are formed, form a finite basis. Complex concepts are learned, that is, they acquired from primitive concepts by means of some computational processes.\footnote{Fodor follows the Empiricist tradition (at least assuming that his interpretation of this tradition is correct) in claiming that concept learning is a technical term referring to a computational process of constructing and confirming hypotheses about objects falling under the concept. More on this below.} Primitive concepts cannot be learned—they have to be acquired in a different way.

\textbf{v. Concepts are public.}

It must be possible for concepts to be shared because otherwise communication would not be possible (cf. Fodor, 1998a, p. 28). Even though concepts, ontologically speaking, are mental \textit{particulars}, they are also (because of their function) symbols, and as such they satisfy a type/token relation (Fodor, 1998a, p. 3). Two people can share the same concept when they have a token of the same concept type. Publicity is also required for explaining intentional behavior by appeal to the person's mental states which are
composed of concepts (cf. Prinz, 2004, pp. 14–5): different people's actions can be motivated by the same attitudes if these attitudes are composed of the same concepts.

4.3 Concept acquisition

Fodor in his 1975 and 1981b is a radical 'mad-dog' concept nativist\textsuperscript{148} in that he claims that all (or almost all) of our lexical concepts are innate.\textsuperscript{149} But how in the first place do we know whether any concepts are innate or not? Well, we need to look at how we end up having them. According to Fodor (in his 1975, 1981b and 1998a),\textsuperscript{150} there are two possibilities here: a concept is either innate, or it is acquired from experience, that is, it is learned. Fodor believes to have shown that the Empiricist account of concept acquisition, according to which (almost) all concepts are learned from experience, is seriously deficient. And since the empiricist claim that most of our concepts are learned from experience is wrong, we must conclude that in fact most of our concepts are innate. Here is how the story goes according to Fodor.

\textsuperscript{148} The term first appears in Fodor's 1984, p. 39.
\textsuperscript{149} In his 1998a, Fodor admits that primitive concepts do not have to be innate. In his 2008, on the other hand, he seems to be saying both that there are no innate concepts at all, and that most of our concepts are innate. I do not think that Fodor is contradicting himself. Rather, while he does reject his earlier definition of the term 'innate,' he still hopes to find a definition that would work. See below 4.3.4 through 4.3.6.
\textsuperscript{150} Fodor's 1975, 1981b and 1998a are the main sources that I use to present his views on concept acquisition.
4.3.1 Concept learning

Empiricists, as Fodor explains in his 1981b, hold that all complex concepts, whether phrasal or lexical, are acquired in a similar way—they are constructed from primitive concepts. This happens, on Fodor's reading of the Empiricist account, in a way similar to the way in which we acquire our native (natural) language\textsuperscript{151}, that is, by means of the process of 'concept learning,' i.e., of an inductive process of hypothesis formation and testing. Here is how Fodor describes this process:

Q116. Concept learning involves the inductive fixation of certain beliefs, and 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, 1981b, p. 267; cf. also Fodor, 1975, p. 42)

Q117. 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, 1981b, p. 268)

Concept learning is an epistemic process: it involves formation of beliefs and the search for evidence. This is what, on Fodor's understanding, abstraction is supposed to mean—where abstraction is the term used by the empiricists to refer to the process of concept acquisition. According to Fodor, whether we talk of hypothesis testing or of abstraction, what we have in mind are "the same inductivist accounts of concept learning" (Fodor, 2008, p. 136, footnote 6). He explains:

\textsuperscript{151} Fodor believes that "learning a (first) language involves formulating and confirming hypotheses about the semantic properties of its predicates," see Fodor, 1975, p. 79.
Q118. It's sometimes supposed that concepts are learned 'by abstraction' from experience with their instances. On this sort of model, experiences of the instances provide evidence about which of the shared properties of things in a concept's extension are 'criterial' for being in the concept's extension. Accordingly, concept learning involves inducing the criteria from samples of the extension. (Fodor, 2008, p. 135, footnote 6)

Concept learning, as Fodor understands it, is of no use if what we try to explain is acquisition of primitive lexical concepts. Primitive concepts have no internal structure, i.e., they cannot be decomposed into constituent features. If we wanted to form and test any hypotheses, they would have to be hypotheses using the primitive concept in question. And so, this way we would not explain how the concept was acquired. In what he calls the Standard Argument (see below), Fodor shows that concept learning is problematic even with respect to complex concepts. It turns out that by means of this method it is impossible to learn any new concept at all. And then, if (almost) no concepts can be learned, radical concept nativism seems to be the only option.

4.3.2 Standard Argument: No learning of new concepts

Here is the gist of Fodor's Standard Argument. 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

153 What Fodor calls concept learning could more plausibly be understood as a process of learning the meaning of words: it makes sense to say that a child learns that the word 'cat' expresses the concept CAT by forming hypotheses about cats; "by observing what people say and do in the presence of cats and other animals" (Davis, 2003, p. 454).
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 immediately obvious that you were able to complete the task only because you already had the two concepts, GREEN and TRIANGLE, at your disposal. If this is really the case, then you did not learn any new primitive concept (even though, supposedly, you did learn something). The concepts you used to figure out what a flurg is were already there in your conceptual apparatus. The conclusion here is that 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, what is called 'concept learning' does not really refer to any process by means of which we would learn any new primitive concepts. The following passage summarizes Fodor's Standard Argument:

The No Learning of New Concepts Argument (NLNC)

Q119. The hypothesis whose acceptance is necessary and sufficient of learning [the primitive 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 the learning of a new concept (viz., a concept distinct from Φ). (Fodor, 1975, p. 95)

Fodor's Standard Argument is supposed to show that the empiricist account of concept learning cannot be applied to the acquisition of primitive concepts. Primitive concepts
which lack internal structure have no components that are semantically interpretable, and so even though they can compose and form more complex concepts, they certainly can not \textit{decompose}. Because they do not have any components, they cannot be derived from any other concept, and so, they cannot be learned. Hypotheses can only be formed using some preexisting concepts. So, for example, assuming that RED is a simple concept, you need some preexisting concepts in order to learn it. But the only concept that would make it possible for you to acquire RED is RED itself. Similarly, 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, 1981b, pp. 269 and 272). So, if primitive concepts cannot be learned, they must be innate.\footnote{In his 1975 and 1998a, Fodor believes that by means of 'concept learning' it is possible to acquire complex concepts. And so, even though a primitive concept RED cannot be learned because it is primitive, the concept FLURG (which means: GREEN AND 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 \textit{LOT2}, Fodor notices that the presented above \textit{No Learning 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. As we have seen, the NLNC Argument 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 \(\Phi\) for some or other concept \(\Phi\)." If this is true, then it 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). If, however, Fodor still 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 Fodor's view even more implausible.}

Based on his discussion regarding the process of 'concept learning' Fodor draws several conclusions about similarities and differences between Empiricists and Nativists.
First of all, he notices that there is a great number of claims that they share. Most importantly, Fodor explains, they both distinguish primitive and complex concepts, and agree that complex concepts are obtained from primitive concepts by means of some combinatorial apparatus. They also agree that acquisition of primitive concepts requires the activation of the sensorium, and that these concepts are unlearned. The main difference between the two positions, according to Fodor, concerns the structure of lexical concepts (Fodor, 1981b, pp. 278–9).

Nativists, Fodor explains, believe that all or most lexical concepts have no internal structure, and are unlearned. For empiricists, only sensory concepts are not learned. This implies that empiricists need to assume that most of our lexical concepts are structured. For a typical empiricist, only a small set of concepts can be directly produced by the sensorium—concepts of colors, shapes, tastes, etc. All other concepts, for empiricists, have to be complex, and acquired by concept learning.

In order to show that the nativists' position is correct and to arrive at the conclusion that most of our concepts are innate Fodor now needs to provide support for two claims:

---

155 For similarities between Empiricism and Nativism according to Fodor, see his 1981b, pp. 274–5 and 277.
156 What Fodor considers here is only one empiricist picture according to which the source of our primitive concepts is sensation. It is not clear, however, that the empiricists have to limit themselves to sensation as the source of concepts. If they allow that we get the concept RED as a result of sensing red objects, why couldn't they also hold that we get the concept CAT as a result of perceiving cats? It would be an advantage for the empiricists to focus on perception as the source of concepts, because it would be easier to explain that we get the concept DOG from an organized percept of a dog, without having to trace it all to the individual sensation concepts.
157 Empiricists support a cognitivist account of concept possession, according to which to have a concept means to know something (cf. Fodor, 1998a, p. 124). This is another reason why they need an inductivist account of concept acquisition.
A. that most of our lexical concepts are primitive in the sense of unstructured, and so, that most of those concepts cannot be learned, and
B. that 'unlearned' implies 'innate.'

Let us first briefly summarize the reasons that Fodor gives for his Informational Atomism, that is, the view that most of our lexical concepts are unstructured.

4.3.3 Informational atomism

According to the theory of concepts that Fodor supports, informational atomism, most of our lexical concepts are primitive: they are atoms, that is, unstructured symbols of LOT, which carry information about features of the world. Why is it that Fodor believes that among our lexical concepts "there are (practically) no complex mental representations at all" (Fodor, 1998a, p. 43)? For a concept to be complex it must have a structure (it must have semantically interpretable components). There are two ways, according to Fodor, in which this may occur: a concept is complex either when it is definable, or if it can be understood as a prototype. As Fodor will try to convince the reader, there are (almost) no lexical concepts with a definitional structure, and, even though some lexical concepts have prototypes, still, the prototype theory of concepts has to be rejected for other reasons. Let me explain.
i. Definitions

According to the classical theory of concepts, complex lexical concepts have internal definitional structures. Every lexical concept that is complex, is composed of simpler concepts that represent the properties which constitute the necessary and sufficient conditions for falling under this concept. That is to say, every complex lexical concept can be decomposed into its constituent concepts which represent the defining properties of things that fall under the concept (necessary and sufficient conditions for falling under the complex concept can be stated in terms of its component concepts). A thing falls under a complex lexical concept if and only if it falls under its component concepts. So, for instance, the concept BACHELOR is complex because it is constituted by two simpler concepts: UNMARRIED and MAN. A thing falls under the concept BACHELOR if and only if it is both 'unmarried' and 'man.'

Even though there are certain clear advantages of the classical theory of concepts (it provides a plausible account of concept acquisition, concept possession, categorization, and reference determination; cf. Margolis & Lawrence, 2005), there are unfortunately too few examples of successful definitional analysis. Most lexical concepts, as Fodor often emphasizes, are non-definable, they do not have a definitional structure, and so, in that sense they are not complex.
ii. Prototypes

A possibility for concepts to be complex without being definable is offered by the prototype theory\textsuperscript{158}. On this non-classical theory of concepts, concepts have probabilistic, or prototype structures. Complex concepts are not defined by a set of necessary and sufficient properties. Rather, instances of a concept are defined by their resemblance to a \textit{prototype}, that is, to a most typical example (or to the best exemplar) of the concept. A prototype, a structured mental representation, specifies the characteristic (and not defining) features (that is, features that are diagnostic, statistically frequent, or salient, but often contingent for category membership; see Prinz, 2004, p. 52) of things that belong to its extension. Whether some thing falls under a concept can be determined by figuring out how similar this thing is to the concept's prototype. (This is done by applying some sort of 'similarity metric.' For instance, the shared features may be simply summed up to see whether a given object resembles an exemplar to a sufficient degree. In more complex cases, a different similarity value of different features may have to be taken into consideration; cf. Prinz, 2004, pp. 53–4.) A prototype shares the maximum number of features with other instances of a given concept and a minimum number of features with instances of other concepts. Different things that fall under the same concept may vary in their degree of family resemblance to the prototypes. For this

\textsuperscript{158} The beginnings of the prototype theory can be traced to Wittgenstein who considered the example of the concept GAME and noticed that it is impossible to determine necessary and sufficient attributes of all things that fall under that concept. There are no defining properties that all games possess. Rather, different instances of games are similar by \textit{family resemblance} (cf. Wittgenstein, 1953, pp. 66–7), without any characteristic being shared in common by all members of the family. See also Rosch and Mervis, 1975.
reason, category membership is a matter of degree. Concepts, therefore, have indistinct boundaries; they may be represented by fuzzy sets.

On the prototype theory, to learn a concept means to construct a prototype out of more basic concepts. According to Fodor, we can learn new concepts using this method, because it is possible to construct a prototype without having the target concept. For instance, based on various experiences with dogs one may construct a prototype: something which has fur, four legs, which has a tendency to bark and wag its tail, and to chase after cats, etc., without previously having the concept DOG. (To have the concept DOG would mean, supposedly, to have encoded in one's head a description of a prototypical dog. It is not entirely clear, however, what would be the difference between having the concept DOG and having a dog-prototype without having the concept.159)

The prototype theory does not face the problem that was fatal for the classical theory of concepts. Fodor admits that, while "definitions are practically nowhere," "prototypes are practically everywhere" (Fodor, 1998a, p. 93). In addition, because, according to Fodor, a concept with prototype structure is still complex, it can be learned even though it is not

159 The distinction between having the concept X and having the X-prototype seems to be crucial for Fodor in his 2008. In this work Fodor suggests that perhaps we could hold that having prototypes is a first stage in concept acquisition (Fodor, 2008, p. 150). On this proposal, we would first get an X-prototype, and only then acquire the concept X. We would explain the process leading to prototype-formation as 'a process of statistical inference,' and the process leading from a prototype to a concept as 'some reliable but not intentional (and hence, a fortiori, not inferential) neurological process' (Fodor, 2008, p. 162). The advantage of this suggestion is, according to Fodor, that it would help us to account for the d/D phenomenon (see below, 4.3.5). Unfortunately, Fodor's proposal faces serious problems. For instance, it is certainly possible to have a concept Y, for instance, without having the Y-prototype at all, or while having something like an anti-Y-prototype. Suppose a child grew up in an animal shelter where all the dogs had three legs, didn't bark, and had no fur. The child could still have the concept DOG but it would be based on an anti-DOG-prototype. Or imagine a different child who grew up with cats, seals, and other animals, but had never seen nor heard of dogs. In a flight of pure imagination this person imagines a barking animal with four legs, fur, etc. Again, it seems that the person would thus have the concept DOG but without having the DOG-prototype at all (I owe both examples to prof. Wayne Davis).
definable (cf. Fodor, 1981b, pp. 292 and 284). Fodor's optimism regarding the prototype theory seems, however, unfounded. Fodor's suggestion that concepts as prototypes are undefinable but still complex is far from obvious. If we want to hold on to the view that concepts are thought-parts, then an undefinable concept is really an undefinable thought-part. If so, then it seems that concepts as prototypes could not be learned. Either we have the thought-part DOG, or not. If you don't have this thought-part, then you cannot even think about dogs. If DOG is undefinable, then you cannot form any hypotheses about dogs unless you have the concept DOG. So, why does Fodor think that concepts as prototypes are undefinable and still complex? Perhaps he could defend his position if he also claimed that the concept DOG is exactly the same thing as the concept PROTOTYPICAL DOG. The latter, presumably, can be learned by forming hypotheses about an object's being furry, barking, wagging its tail, etc. But DOG and PROTOTYPICAL DOG are certainly not the same; they certainly do not have the same extensions.

Fodor does not consider this problem. (In fact, he returns to the idea of acquiring prototypes by means of hypothesis testing method in his 2008.) Instead, he points out to a different problem with prototypes: the prototype theory of concepts does not satisfy one of the necessary features of concepts (listed above in 4.2), the requirement for compositionality. Prototypes just do not compose. This can be seen for instance in the fact that the prototypes of complex concepts often have little or nothing in common with the prototypes of their component concepts (cf. Fodor, 1998a (p. 100) where Fodor
contrasts the prototype of PETFISH with the prototypes of FISH and PET. Since concepts do compose, they cannot be understood as prototypes.

If Fodor is right in his rejection of concepts as definitions and concepts as prototypes (and in the assumption that these are the only ways in which concepts could be complex), then most of our concepts that are not phrasal, are primitive, non-complex, and cannot be learned. The question we need to ask now is whether all those primitive concepts really must be innate.

### 4.3.4 Does 'Unlearned' imply 'Innate'?

In order to figure out Fodor's answer to the question of why 'unlearned' means 'innate' we need to look at how he explains in more detail the empiricist position on concept acquisition. On Fodor's interpretation of the empiricist view, as we have seen, complex concepts are acquired by means of the inductive process of concept learning. The empiricist idea of concept learning on its own requires, without doubt, quite a lot of innate stuff. If we acquire any concept 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, 1981b, pp. 268–9 for what such a machine must include). So, it seems common sense that for empiricists, complex concepts are learned by means of an innate mechanism of concept learning. The empiricists also hold that in addition to complex concepts, there is a small set of primitive sensory concepts which are not learned, but whose "acquisition is occasioned by the activation of the
sensorium" (Fodor, 1981b, p. 272). And this, according to Fodor, is enough to show that for the empiricists sensory concepts are innate. Let me explain.

A crucial and distinguishing characteristic of concept learning is a rational connection between the concept and what it represents (Fodor, 1981b, p. 275). The concept FLURG (or BACHELOR, etc.) is complex. The way in which it is acquired is rational-causal in the sense that "the experiences which eventuate in the availability of such a concept are held to bear a confirmation relation to some hypothesis which specifies the internal structure of the concept" (Fodor, 1981b, p. 272; cf. also Fodor, 1981b, p. 279: "Experiences bestow inductive warrant upon hypotheses which articulate the internal structure of the concepts."). In the case of those concepts which are learned, i.e., acquired in a rational-causal way, we can provide an intentional description of the acquisition process. We can show that the acquisition of the concept is rationally related to the experiences that gave rise to it, or that it 'makes sense' in the light of those experiences.

On the other hand, in the case of primitive concepts, such as RED, or DOG, or, plausibly, DOORKNOB and SPATULA, there is no rational connection between the concept and the experience that triggers it. The experiences of the organism do not stand in a confirmation relation to the concepts whose availability they occasion (Fodor, 1981b, p. 273); rather, they "function as the innately specified triggers of the concepts which they (...) release" (Fodor, 1981b, p. 280). We acquire concept RED not by learning, but by "opening one's eyes and looking" (Fodor, 1975, p. 97). All that is needed for us to acquire RED is a properly functioning sensorium, and the right kind of stimulus. For
this reason, the only kind of explanation we can have of why certain experiences lead to certain concepts is that each sensory organ is a mechanism which realizes "a function from stimuli to primitive concepts" (Fodor, 1981b, p. 265), and that "the structure of the sensorium is such that certain inputs trigger the availability of certain concepts" (Fodor, 1981b, p. 273). The connection between a primitive concept and the experience that triggers it is, therefore, brute-causal.

Fodor warns us not to be fooled into thinking that what we have here is an explanation of how primitive concepts are acquired from experience. The structure of the sensorium (together with the rules of computation) does determine what concepts are potentially available to an organism (Fodor, 1981b, p. 265; cf. below, 5.3.3). However, Fodor claims, it does not provide us with any explanation of how such a mechanism works.

All right, so we have here a contrast between concepts that are learned from experience, and those that are merely triggered by experience. Even though talking about sensorium does not immediately provide us with an explanation of how sensory concepts are acquired, still, it is far from obvious that the mechanism which allows us to obtain these concepts also requires that concepts themselves be innate. As it was the case with concept learning, all which certainly needs to be innate here is the mechanism that realizes the function from sensory stimuli to sensory concepts. Whether concepts as well have to be innate can be really determined, it seems to me, once we are clear on what exactly we mean by the term innate (this will be discussed later in chapter V) and also after we provide a careful analysis of how the mind functions, or, in other words, of the
account of the architecture of the mind that we'd want to accept. Fodor himself does not take that route. The answers he gives, however, are far from satisfactory. Here are some ways in which we can try to explain what still provides some kind of support for Fodor's mad-dog nativism.

i. What would support Fodor's mad-dog nativism?

First of all, in his conclusion Fodor seems to follow Descartes. Descartes explains that 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

Q120. ideas [or concepts] of movements and figures are themselves innate in us. So much the more must be 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").

What we see in the passage quoted is a very quick switch from 'innate faculty' or mechanism to innate ideas. Fodor does not seem to think that such a switch requires any kind of justification. Instead, he appears to assume that the inference from 'unlearned' to 'innate' is equally obvious.

---

Another possibility is that it is all just a terminological matter. Perhaps for Fodor there are only these two options: either there is a rational connection between a concept and what it represents, or the concept must be innate. If so, then indeed, anything that is not learned must be innate. Similarly, Fodor seems to simply define 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, 1998a, p. 129). Unfortunately, he does not provide any satisfactory explanation of what triggering is really supposed to be, and why it requires innate concepts.

Fodor has recently considered some of the conclusions I have reached, for some of the same reasons. In his LOT2, we read him saying that we "can't infer from a concept's not being learned to its being innate" (Fodor, 2008, p. 144), because "'learned' and 'innate' don't exhaust the options" (Fodor, 2008, p. 130). He admits that "it mustn't be taken for granted that learning is the only way that a concept might be acquired" (Fodor, 2008, p. 132, footnote 2). It is even less clear than before, however, what meaning these terms are supposed to have in the end, because at a certain point Fodor seems to be left with the unsatisfactory claim that concepts are neither learned nor innate. Fodor does not explain what alternative we are left with. I try to provide an answer to this question in chapter V.

ii. Poverty of Stimulus

The best reason that Fodor gives in support of the claim that primitive concepts must be innate is based on arbitrariness of the triggering process. According to Fodor, the crucial
feature of the triggering relation between lexical concepts and their occasioning experiences is that it may be extremely random (I'll talk about it more in detail below; cf. also Fodor, 1980, p. 280). This means that looking at the experience that triggered a concept does not explain why exactly this concept, and not a different one, was acquired. And so, when we say that primitive concepts are triggered by experiences, we hardly explain anything. In fact, as Fodor emphasizes, the environment provides a poor basis for the concepts that the organism acquires (Fodor, 1981b, p. 280). The data sample from the environment is 'fragmentary and impoverished,' and does not seem sufficient to explain why a given concept is acquired. And so, Fodor believes, we should conclude that "the concept isn't coming from the environment, it's coming from the organism" (Fodor, 1981b, p. 280).

The poverty of stimulus argument seems to offer a stronger support to the thesis that 'triggered' implies 'innate.' If the environment does not provide a sufficient explanation for how we acquire concepts, then indeed we might want to look for an explanation on the side of the organism. Still, it is not obvious that the innate elements we will find on the side of the organism are indeed concepts.

4.3.5 Problems with triggering

i. The d/D problem

Fodor notices that the idea that primitive concepts are brute-causally triggered faces a serious problem. As we have seen, 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).

Acquiring a concept in a brute-causal way is similar, Fodor believes, to acquiring it as a result of being hit on the head. But this kind of concept acquisition seems completely random. And if it is random, then we have no way to explain why it is experience with doorknobs and not giraffes that causes the concept DOORKNOB to be acquired. This is what Fodor calls the 'd/D problem': the triggering relation does not provide any way to explain the connection between the concept produced and the object perceived.

Fodor rejects the idea of concepts being innate in the sense of having their content independently of any encounter with the environment (cf. Fodor, 1981b, pp. 274–5). Rather, he is looking for a way to show why interaction with things of a certain sort 'triggers' them into having the contents that they actually do have. And so, he wants a theory which will be able to explain why there is what he calls "a content relation" between the experience that eventuates in concept attainment and the concept that the experience eventuates in attaining" (Fodor, 1998a, p. 132).

Doorknobs are certainly the best source of evidence for DOORKNOBS, but if we talk of 'evidence,' then it seems that some kind of learning must be involved in acquiring the concept DOORKNOB. According to Fodor, it is clear that what could solve the d/D problem is the hypothesis testing method, a method which does guarantee that there is a

---

161 Fodor talks about the "relation between the content of the concept and the content of the experience that occasions it" (Fodor, 1998a, p. 130). What he really seems to point out to is the need to explain the fact that in concept acquisition there is a match between the content of the concept (e.g., the concept RED), and the experience that occasions it (e.g., experiencing something as red (Fodor, 1998a, p. 131).
reliable relationship between experiences and concepts that they cause. Unfortunately, as we have seen, this method is of no use when we want to explain acquisition of primitive concepts. In addition, we may question whether hypothesis testing method really does solve the d/D problem. Fodor points out that it remains unexplained why we get the concept DOORKNOB from experiences of doorknobs, and not from experiences of cats. And it seems that he should worry as well about the question why we form hypotheses about doorknobs, when perceiving doorknobs, and not when perceiving cats.

Fodor, however, never considers this issue. He seems to believe that it is unproblematic and obvious: when we have experiences of cats, we don't form hypotheses containing DOORKNOB, hypotheses about doorknobs. This is, he believes, because experiences with doorknobs constitute evidence for our hypotheses. In the case of brute-causal acquisition of concepts, there is no place for evidence; that's why here there is a need for an explanation of the match between the concept and the experience. Still, if Fodor does think that it is mysterious that we don't get the concept CAT by looking at a doorknob, it seems that he should consider as equally mysterious the fact that our experiences with cats serve as evidence for our CAT concepts, and not for the concept DOORKNOB.

Because Fodor believes that the hypothesis testing method solves the d/D problem he really wants to hold on to it. Since he realized that hypothesis testing does not work as an explanation of the acquisition of concepts, in his 2008 he suggests that perhaps it would work for the acquisition of prototypes. Acquisition of prototypes would then be the first stage in the process of concept acquisition. And because prototypes would be learned by a rational inductive process, they would fit the world—they would not face the d/D problem. Unfortunately, this solution only pushes back the d/D problem. The question that now remains mysterious is why a prototype D leads to the acquisition of the concept D, instead of the concept F. (Fodor's imaginary interlocutor in his 2008, Snark (who replaced Granny from Fodor's earlier works) does notice this problem, in Fodor, 2008, p. 165.)
ii. Could Evolution Help?

Fodor considers another way in which we could try to deal with the d/D problem. It would be to suggest that it is evolution that guarantees that there is a connection between our experiences and concepts that get triggered by them. And so we might wonder:

Q121. Wouldn't a mechanism that triggers the concept X consequent upon experience with Xs be more of a help with surviving (or getting reproduced, or whatever) than, say, a mechanism that triggers the concept X consequent upon encounters with things that aren't Xs? (Fodor, 1998a, p. 128)

If so, if we accept this kind of reference to evolution, then this "might explain why the relation between the content of experiences and the content of the concepts that eventuate in locking to it is so rarely arbitrary" (Fodor, 1998a, p. 128).

According to Fodor, there are problems with this suggestion. He believes that even if we managed to provide a plausible account of how evolution equipped us with reliable triggering mechanisms which would guarantee that "it is instances of F-ness (and not of G-ness) that trigger the concept F" (Fodor, 1998a, p. 129), this would still take us back to extreme nativism because "you can only trigger what's already there" (Fodor, 1998a, p. 129).

And so, Darwin is anyway of no help to us at all, because, according to Fodor, we could never get from evolution what we want from it:

Q122. Even if the doorknob/DOORKNOB relation is selected for by evolution, what, if not inductive learning could be the mechanism by which it is implemented? If concept
acquisition isn't inductive, then how does Mother Nature contrive to insure that it is instances of F-ness (and not of G-ness) that trigger the concept F in the course of ontogeny? After all, if Mother N wants to select for the doorknob/DOORKNOB type of relation between concepts and their experiential causes, she has to do so by selecting a mechanism that produces that relation between one's concepts and their causes. (...) The obvious candidate to select if one wants to ensure that concept acquisition exhibits the d/D relation is inductive learning. (Fodor, 1998a, p. 129)

Fodor holds that the only mechanism that could possibly help, i.e., that would guarantee the needed connection between experiences and concepts, is the inductive mechanism of hypothesis formation and testing. And this one, yet again, is useless with respect to primitive concepts.¹⁶³

¹⁶³ Fodor also tries to solve the d/D problem with what he calls the Constitution Hypothesis. If hypothesis testing worked for primitive concepts, it would solve the d/D problem, Fodor tells us, in virtue of providing an epistemological explanation about how concepts are acquired. Instead, the solution offered by Fodor is "an explanation which makes the d/D effect the consequence of a metaphysical truth about how concepts are constituted" (Fodor, 1998a, p. 133). His idea is that we can get some help if we look at concepts of secondary qualities (Fodor, 1998a, pp. 130–1). Such concepts are also susceptible to the d/D problem. Concept RED, because it is primitive, must be acquired in a brute-causal way. And it seems completely random and unexplained that it is red things that trigger in us the concept RED. They could as well trigger the concept BLUE, or DOORKNOB. However, as Fodor notices, "the relation between the content of a sensory concept and the character of its cause is not arbitrary when the cause is intentionally described" (Fodor, 1998a, p. 130). We realize that it is simply the experience of something as red, or the fact that something appears red to the subject, that causes the occurrence of the concept RED. And so, the d/D problem is solved in the case of secondary qualities concepts once we focus on the fact that these concepts refer to appearance properties. This is because the point about appearance properties is that they don't raise the question (...) 'What is it that the things we take to be Xs have in common, over and above our taking them to be Xs?'" (Fodor, 1998a, p. 135). (Contrary to what the above quote may suggest, Fodor does not really equate 'taking something to be X' with 'experiencing something as X.' Rather, what he wants to claim is that this is typically the case with appearance properties. And so, Fodor would agree that we can take something to be X through all sorts of belief-forming operations, not necessarily as a result of experiencing it as X. Also, we can take something to be X without experiencing it to be X.) And so, Fodor suggests, "how come DOORKNOB is generally learned from doorknobs? is to be answered in the same way that Locke dealt with 'How come it's typically red-sensations that red things cause us to have?' The answer, in both cases, is 'that's of the essence of the properties concerned'" (Fodor, 2001, pp. 143–4). Fodor's suggestion is, therefore, that perhaps also "doorknob works like 'red'" (Fodor, 1998a, p. 135), and that really "there are lots of appearance concepts that aren't sensory concepts" (Fodor, 1998a, p. 135). If we accept this view, then we would conclude that "what it is to be a doorknob isn't evidenced by the kind of experience that leads to acquiring the concept DOORKNOB" (Fodor, 1998a, p. 134) but it is constituted by this experience. On this proposal, doorknobs are mind-dependent entities, that is, they are the types of thing that cause minds like ours to acquire DOORKNOB. And to have the concept DOORKNOB means to 'resonate' to the property that this concept expresses, i.e., to 'doorknobhood' (Fodor, 1998a, p. 137). This way we reach a satisfactory solution to the d/D problem, because 'If being a doorknob is a property that's constituted by how things strike us, then the intrinsic
I do not think that it is right to dismiss evolution so quickly. After all, given that the term 'innate' originated in biology, it may be that those accounts of innateness that in some way refer to evolution end up being more plausible. If Fodor were right that this just takes us back anyway to extreme concept nativism, then we would have to deal with several quite problematic issues. For instance, we would have to ask why evolution would equip us with such innate concept as DOORKNOB. And also, it would not be clear how we'd answer the question how evolution could ensure that doorknobs would trigger the innate concept of a doorknob rather than the innate concept of a giraffe. But I don't think that Fodor is right. If we do refer to evolution, we may find a satisfactory solution to the d/D problem. Also, the conclusion will be that extreme concept nativism should in fact be rejected. These issues will be considered below in chapter V (see especially 5.4.4).

4.3.6 Are concepts really innate?

Independently of whether we manage to solve the d/D problem, the question that we still need to consider is how plausible it is, on Fodor's account of cognition, to hold that

connection between the content of DOORKNOB and the content of our doorknob-experiences is metaphysically necessary" (Fodor, 1998a, p. 136). If the connection is metaphysically necessary, then obviously it is no longer arbitrary. It becomes metaphysically necessary that the d/D phenomenon occurs, that is, that doorknobs cause us to acquire DOORKNOB, similarly as it is in cases when we get the concept RED when perceiving red things, where it is metaphysically necessary that experiences of red (things) cause us to acquire the concept RED. For lack of space I won't consider the plausibility (or lack thereof) of Fodor's metaphysical solution to the d/D problem.
concepts are innate. In his 1998a, Fodor himself notices at some point that on the account he proposes

Q123. 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. (…)

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 of anything that has intentional content. Since the sensorium isn't an idea, it is a fortiori 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, 1998a, p. 142)

Fodor does not like this conclusion. Everybody agrees that sensorium is innate, so the conclusion that we seem to have arrived at is just a boring claim.

Notice, however, that it is not clear why Fodor arrives at such a conclusion in the first place. After all, he still holds that we acquire the concept DOORKNOB in a brute-causal way. That is, he'd say that our concept DOORKNOB must be triggered by doorknobs. But we have seen that for Fodor we can only trigger 'what's already there, waiting to be triggered'; and so, we are back to mad-dog concept nativism.

At this point we need to look more closely at Fodor's account of the architecture of the mind and see what it implies with respect to the issue of concept acquisition and the supposed innateness of concepts. As we said, concepts, for Fodor, are mental symbols that the mind manipulates in its cognitive processes. Once we have a clearer account of how the mind actually functions, it should become easier to figure out how it acquires the concepts that it manipulates.
4.4 Fodor's views on the architecture of the mind

4.4.1 Three types of mental mechanisms

The mind, according to Fodor, has three components/kinds of cognitive systems, transducers, input systems and central processors\textsuperscript{164}. Transducers link the cognizing subject to the external world. Input systems are responsible for the analysis of the input provided by transducers. The main role of central processors is the fixation of belief (Fodor, 1983, p. 112). The three kinds of cognitive systems that Fodor mentions correspond quite closely, it seems, to Aquinas's three levels of cognition in the following way: transducers correspond to Aquinas's external senses, input systems—to Aquinas's internal senses together with the intellect's first operation, i.e., what I call 'thinking,' and the central processors correspond to Aquinas's higher intellectual faculties.

If I am right in my claim that Fodor's and Aquinas's accounts of the architecture of the mind are sufficiently similar, then, assuming also that we can find a definition of innateness that both philosophers could accept, we'll conclude that they should take a similar position with respect to the empiricism-nativism debate about concepts.

\textsuperscript{164} Fodor admits that these categories are not supposed to "exhaust the types of psychological mechanisms that a theory of cognition might have reason to postulate" (Fodor, 1983, p. 42).
i. Transducers

Transducers, according to Fodor, are those mental mechanisms that link the cognizing subject to the external world. Their role is to obtain information about how the world is, and also, to allow the subject to satisfy its desires by appropriate acting in the world.

Input transducers (e.g., the retina) acquire information from the external environment and transform it into the format that can be understood and used by input systems\(^{165}\); they transform stimulations from the external world, i.e., physical, non-symbolic input such as light, sound, scent, etc., into neural signals (some patterns of nerve impulses in the brain, or some pieces of brain-code) — symbols that can be understood by input systems (Fodor, 1983, p. 41). Transducers are subsidiary systems (Fodor, 1983, p. 39) which affect the computations of input systems "in ways that are responsive to the flow of environmental events. The function of these subsidiary systems is to provide the central machine with information about the world; information expressed by mental symbols in whatever format cognitive processes demand of the representations that they apply to" (Fodor, 1983, p. 39).

The role of transducers is to specify "the distribution of stimulations at the 'surfaces' of the organism" (Fodor, 1983, p. 42). There are two kinds of stimuli. Patterns of energy that affect the receptors of the organism are called proximal stimuli. And patterns of energy that emanate from objects are called distal stimuli. A proximal stimulus is a disturbance that happens at the surface of the body and is directly

\(^{165}\) There are also output transducers which take symbols as input, and produce non-symbolic output, e.g., neural firing causing a muscle contraction. I won't consider transducers of this kind in what follows.
responsible for the reaction of the receptor. Distal stimuli emanate from 'local distal objects,' that is, from things as they actually exist in the world (e.g., dogs or doorknobs). Proximal stimulus is caused by objects that are a bit further from the body, not immediately next to its surface: that's why these objects are 'distal.' They are also 'local,' because they must be in the vicinity of the cognizing subject—otherwise transducers would not receive any information from them.

The information registered by the receptors is called 'proximal effect,' or 'representation of proximal stimulus' (cf. Fodor, 1985, p. 2; a change in the state of the retina would be an example of such a proximal effect). Input systems abstract from neural signals generated in response to proximal stimuli. The product resulting from such abstractions corresponds to the distal stimulus (in the case of veridical experiences). Fodor explains:

Q124. The character of transducer outputs is determined, in some lawful way, by the character of impinging energy at the transducer surface; and the character of the energy at the transducer surface is itself lawfully determined by the character of the distal layout. Because there are regularities of this latter sort, it is possible to infer properties of the distal layout from corresponding properties of the transducer output. (Fodor, 1983, p. 45)

Thus, we see that 'proximal effects' are directly linked to their 'distal causes' (that is, in our example, to "alterations in the arrangement of the distal objects that radiate and reflect the light," see Fodor, 1985 p. 3). And this is what allows certain 'higher' cognitive mechanisms that have access to proximal effects to infer, on their basis, the arrangement of local distal objects.
Input transducers, for Fodor, stand for the sensorium. The senses are "receptive mechanisms that satisfy the technical notion of a transducer" (Fodor, 1981b, p. 264). "They convert the particular form of energy to which each is attuned into the electrical energy of the nerve impulse" (Fodor 1975, p. 46, footnote 14). (In his 1981b, Fodor describes the sensorium as "the realization of a function from stimuli onto primitive concepts" (Fodor, 1981b, p. 265). Clearly, however, transducers by themselves cannot play this role. On the contrary, input systems are needed for concept production. Therefore, in Fodor's terminology, the sensorium may sometimes include both transducers and input systems. This is analogous to Aquinas for whom the sensorium includes both the external and the internal senses.)

ii. Input systems

Input systems lie between transducers and central systems. Input systems are different from transducers mainly because of the following features:

(1) Transducers operate in an automatic way, and not by means of computation; their job is rather simple. On the other hand, input systems are computational mechanisms which perform complicated, inference-like transformations. Although Fodor believes that input systems are similar to reflexes in virtue of the fact that they are fast, automatic, and innately specified, he also holds that they are computational. Because of this feature they are closer in their character to the thoughtful processes of decision making and planning.
(2) Transducers, ideally, preserve the informational content of the stimuli that they receive from the external world and only alter "the format in which the information is displayed" (Fodor, 1983, p. 41). On the contrary, input systems do not "translate from the representations that transducers afford" (Fodor, 1983, p. 42). Rather, they "perform quite complicated, inference-like transformations" (Fodor, 1983, p. 41).

(3) While transducers specify "the distribution of stimulations at the 'surfaces' (as it were) of the organism, the input systems deliver representations that are most naturally interpreted as characterizing the arrangement of things in the world" (Fodor, 1983, p. 42).

The most important feature of input systems, according to Fodor, is that they are modular. Generally speaking, modularity of mind is the view according to which

Q125. cognition is subserved by a number of innately channeled, domain-specific systems whose operations are largely independent of, and inaccessible to, the rest of the mind. (Carruthers and Chamberlain, 2000, p. 1)

Fodor defines a module as

Q126. an informationally encapsulated computational system—an inference-making mechanism whose access to background information is constrained by general features of cognitive architecture, hence relatively rigidly and relatively permanently constrained. (Fodor, 1985, p. 3)

---

366 Clearly, it would not be plausible to claim that they preserve all the informational content of the proximal stimuli (for instance, touch receptors don't pick up the sounds an object is making, nor do they detect its magnetic field). Fodor, however, should not have a problem accepting the idea that most likely there is some loss of information at the level of transducers (cf. also Dretske, 1981b, p. 147). When he talks about transducers preserving information he rather has in mind the idea that the role of transducers is just to translate the information that they receive into language that the input systems can understand. Their role is not to produce any additional content.
Modules, according to Fodor, are domain specific. A module can be specialized either in the sense that it produces a certain kind of output, or in the sense that it has access only to a specific kind of information provided by transducers, or in both of these senses together. A domain of a module is "the range of questions for which this processing system has been designed to supply answers" (Botterill and Carruthers, 1999, p. 62). What kinds of domains we should distinguish is an empirical question. It seems very plausible that there are separate modules for instance for language acquisition, face recognition, mind reading, naming abstract things, naming living things, hearing of speech, hearing of music, hearing of environmental sounds, etc. (cf. Botterill and Carruthers, 1999).

The second essential feature of modules is their informational encapsulation (impenetrability). An encapsulated system, Fodor explains, "is allowed to 'look at' only a certain restricted class of data in determining which hypothesis to accept" (Fodor, 1983, p. 110). Modules, according to Fodor, do not have access to the information processed by central processes (nor to the information from other input systems).

The encapsulation of modules is similar to that of reflexes. Consider, for instance, the blink reflex. It doesn't help that you know a person's character and you know that she won't stick her finger into your eye. If she quickly moves her hand too close to your eyes, you will blink. This is because the reflex does not have access to your knowledge of the person's character (Fodor, 1983, p. 71). And so, Fodor explains, one important aspect of encapsulation of modules is "their insensitivity to the utilities of the organism" (Fodor, 1983, p. 102). Encapsulation is very important for survival because it prevents us from
seeing the world the way we wish it was; instead, it assures that we see it the way it really is independently of our wishes. The fact that perception is isolated "from certain effects of background belief" is responsible for both the speed and "the objectivity of perceptual integration" (Fodor, 1983, p. 43). (Fodor says in addition that encapsulation is what explains "the veridicality of perception given that the world doesn't always prove to be the way that we would prefer it to be" (Fodor, 1983, p. 102). But clearly, even though encapsulation may be useful if it prevents wishful thinking, this is not yet enough to guarantee that perception always presents the world the way it really is.)

Modules, according to Fodor, operate automatically and without the intervention of higher cognitive processes. They operate by means of computational processes that are insensitive to the organism's overall belief system (cf. Fodor, 1983, p. 117). In addition, the functioning of input system is not voluntary: it is not a matter of decision that when you say 'I like you,' I understand what you said. Similarly, a human cognizer "can't help seeing a visual array as consisting of objects distributed in three-dimensional space" (Fodor, 1983, p. 53).¹⁶⁷

The role of input systems is to allow the subject to identify objects in the external world. Their input is the symbolic output produced by transducers, and their output is the representation of "the arrangement of things in the world" (Fodor, 1983, p. 42). On the basis of the information provided in the symbolic format by the transducers, input systems—by means of computation—infer how the external world really is. In Fodor's words,

¹⁶⁷ Fodor also discusses other features of modules that I don't present here for lack of space.
Q127. Input 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)

Input systems are "mechanisms for projecting and confirming hypotheses" (Fodor, 1983, p. 110). After the organism's transducing mechanisms register proximal stimuli, input systems get activated, and they infer a hypothesis about distal objects in the organism's external environment. Since the role of inferring the arrangement of local distal objects belongs to perception, input systems, for Fodor, stand for perception plus language. 168

Perception is "a process in which representations of proximal stimuli causally determine beliefs about distal layouts" (Fodor, 1989, p. 4). Specific beliefs about external objects are inferred from representations of proximal stimuli.

Modules are supposed to be 'computational mechanisms,' and their role—to produce representations of distal objects. The only computations that Fodor mentions, however, are inference-like processes, which take premises and produce conclusions. It is not clear how such computations can produce representations of distal stimuli. I'll come back to this issue below (see Conclusion).

168 Fodor acknowledges that "the identification of input processing with perceptual analysis is itself only approximate" (Fodor, 1983, p. 43). He also explains why it is perception and language that are both classified as input systems: "Both serve to get information about the world into a format appropriate for access by such central processes as mediate the fixation of belief" (Fodor, 1983, p. 46). "Just as patterns of visual energy arriving at the retina are correlated, in a complicated but regular way, with certain properties of distal layouts, so too are the patterns of auditory energy that excite the tympanic membrane in speech exchanges. With, of course, this vital difference: What underwrites the correlation between visual stimulations and distal layouts are (roughly) the laws of light reflectance. Whereas, what underwrites the correlation between token utterances and distal layouts is (roughly) a convention of truth-telling. In the root case, the convention is that we say of x that it is F only if x is F. Because that convention holds, it is possible to infer from what one hears said to the way that the world is" (Fodor, 1983, p. 45).
iii. Central processors

Central cognitive systems are responsible for processes of memory, thought, and decision-making. They take conceptual input provided by modules, and generate conceptualized output, that is, information relevant to the subject's belief system. As Fodor explains, central processors are influenced not only by information provided by input systems but also by what the system as a whole already knows. They have access to all cognitive domains of the whole system and to memory. Their role is to make sure that the organism's beliefs are a function of all the information that it acquires and that it already has. Central processors are not domain-specific like modules, but rather domain-neutral; they cut across cognitive domains. They are rather slow,¹⁶⁹ often conscious, non-mandatory; they can be controlled. The main function of central processors is "the fixation of belief (perceptual or otherwise) by nondemonstrative inference" (Fodor, 1983, p. 104).

Fodor compares the work of central processors to Aristotle's 'common sense' but he uses the term in a significantly different way than it was used by Aristotle or Aquinas. Central processors, Fodor suggests, put together and simultaneously consider "the representations delivered by the various input systems," refer to the background knowledge stored in memory, and finally formulate a belief which is, supposedly, the best available "hypothesis about how the world must be given these various sorts of data" (Fodor, 1983, p. 102). When there is a need for it, central systems may correct

¹⁶⁹ According to Fodor, this is, among others, because they are not hardwired, not associated with a fixed neurological structure.
representations of the distal layout provided by a given input system in light of background knowledge of the subject, and of the results obtained from other input systems.\footnote{170}

(For lack of space, in this dissertation I don't go beyond the brief characterization of central processors just offered.)

### 4.4.2 Fodor vs. Aquinas on the architecture of the mind

Even the above rather summarized presentation of Fodor's account of the architecture of the mind is enough to show significant similarities between the views on this topic presented by Fodor, and those by Aquinas. Although Aquinas, for obvious reasons, didn't speak the language of contemporary physics or physiology, it seems quite plausible that if he was aware of the current results of these sciences, he wouldn't have a problem accepting them.

Fodor's transducers correspond to Aquinas's external senses. The role of both is to register stimuli coming from the external world and to translate information that they receive into a code that will be understandable to higher-level cognitive faculties. Also, neither transducers nor the external senses are creative in their functioning.

\footnote{170 Central systems are responsible not only for belief fixation, but also for the production of speech by means of which we "communicate our views on how the world is" (Fodor, 1983, p. 102), that is, we \textit{externalize} our beliefs. So, mechanisms responsible for the production of speech cannot be encapsulated.}
In what concerns input systems, Aquinas certainly didn't talk about anything similar to modularity. Fodor, therefore, would classify him among what he calls horizontal faculty psychologists. According to horizontal faculty psychology,

Q128. cognitive processes exhibit the interaction of such faculties as, e.g., memory, imagination, attention, sensibility, perception, and so forth (...). However, the character of mentation is more or less independent of its subject matter; the faculties are supposed to be invariant from one topic of thought to the next. (Fodor, 1983, p. 11)

Fodor, on the other hand, considers himself a vertical faculty psychologist. On his view, cognitive faculties are domain specific.

What is crucial for our purposes, however, is the role that the input systems play and the fact that they are supposed to be interpreted as innate cognitive mechanisms. Both input systems and Aquinas's internal senses (especially the cогitative power) can be considered as cognitive mechanisms which take as input the output of the transducers/external senses and produce representations of external—distal—objects. They allow the cognitive subject to identify objects in the external world.

---

171 "A horizontal faculty is a functionally distinguishable cognitive system whose operations cross content domains" (Fodor, 1983, p. 13).
CHAPTER V
WHAT IS INNATE IN COGNITION?

5.1 Most common accounts of innateness

The problems with figuring out whether Fodor is or should be a concept nativist (and how much sense his presumed mad-dog nativism really makes) 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 this question 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 what I aim for in this dissertation is not to argue that given his account of the architecture of the mind Fodor should or should not be a concept nativist in some universal sense of 'innate.' As we shall see, there is no such thing. My goal, therefore, is more modest. I only want to show that Fodor should be as much of a concept nativist (or as much of a concept empiricist) as should be Aquinas. Still, I am going to try to get closer to a definition of innateness that is useful in the debate about mental representations and that both philosophers could accept.
First of all, and just for the record, in the present discussion 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, and most importantly, within developmental nativism, it is crucial to point to a distinction between representational nativism (RN) and architectural nativism (AN). Representational nativism (RN) 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 the innateness or pre-specification of the mental representation is explained in different ways (I'll talk about it below). Architectural nativism does not focus on the character or source of our representations which are the output of the cognitive process. 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 organized, at birth, into some innate structures, and that there are some innate ways in which information coming from the external world can 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.\textsuperscript{172}

The present discussion is concerned with the question whether given Fodor's account of concepts and of the architecture of the mind it makes sense for him to hold on to his mad-dog nativism, which is an extreme version of representational innateness.

What we will need to figure out is how representational nativism could be explained in more detail, and in what sense we can talk about innate representations. Before that, however, let us step back and look at various more general explanations of the different meanings of the term 'innate.'

(i) Innate I: Possessed from birth

On the most commonly known version of nativism, a feature is considered to be innate if it is possessed from birth (cf. Locke's \textit{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 specific smells or sounds).

\textsuperscript{172} Cf.: "Locke thought too obvious to mention (…) the existence of natural faculties such as perception, understanding and memory, and innate mental powers like those of abstraction, comparison and discernment. The 'white paper' metaphor is meant to indicate that the understanding (and hence the mind) is originally empty of objects of thought like ideas; but it has whatever apparatus is necessary to acquire them through experience, and then to derive knowledge by comparing and contrasting them with each other" (Harris, 1977, p. 27).
(ii) Innate II: Independent of sensation

'Innate' is also sometimes taken to mean something that is *a priori* to the operation of sensibility, something that is not derived from the operation of sensibility, but from the mind itself, or more generally, something that 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 innate 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, 1981b, p. 280).

This definition of innateness, however, is not very plausible. In fact, if it was accepted, it would turn out not only that there are no innate concepts, but also that there is just *nothing* innate because basically everything in the human development requires at least minimal interaction with the environment, and thus, also sensory experience. It seems plausible that, for instance, walking is an innate capacity. 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 ever able to acquire this innate feature).

The reason why these first two explanations of innateness were popular is that they express a common-sense intuition according to which the term 'innate,' which literally means 'in-born' (cf. its etymology: from Latin prefix *in-* and *natus*, perfect active
participle of nascor, 'born'), stands for something that is 'not acquired' or 'not learned.' One may be tempted to believe that if a trait is not acquired, then, well, we have just always had it, and so, we must have had it at birth. Also, the fact that learning requires the use of our senses may seem to suggest that for something to be innate it means to be acquired independently of sense experience. This, however, would not be a legitimate inference. Even if sensory experience is indeed a necessary condition for learning to take place, and even if 'acquired by learning' is defined as the opposite of 'innate,' all that this would imply is that (a) if something is acquired absolutely independently of sensory experience, then it must be innate, or that (b) if something is not innate, then its acquisition must involve sensory experience. It does not preclude, however, the possibility for something to be innate but still involving or even requiring sensory experience.

(iii) From Not acquired by any psychological process to Triggered

Samuels (Samuels, 2002) tries to avoid the problems of the first two common sense explanations of the term 'innate,' while keeping the intuitions that they stemmed from. He offers what he calls a primitivist definition of innateness according to which a cognitive structure is innate if it is "not acquired by any psychological process or mechanism" (Samuels, 2002, p. 234; see also Cowie, 1999, p. 113). Such a structure is a 'psychological primitive' in the sense that there is "no correct scientific psychological theory that explains [its] acquisition" (Samuels, 2002, p. 246), but there is (or, in principle, could be) a biological theory that explains it. Clearly, a cognitive structure that
is innate in this sense, i.e., whose acquisition cannot be explained by some scientific psychological theory, may not be present at birth. It is possible that it gets acquired later on in the organism's development, and that the non-psychological biological process by means of which it is acquired involves interaction with the environment and some kind of sensory experience.

Fodor seems to have in mind a similar, although more limited, understanding of 'innate' in his 1975, 1981b and 1998a where he takes it to mean 'acquired brute-causally,' or 'triggered.' 'Innate,' in that case, is the only alternative to 'learned,' and 'learned,' for Fodor, means exactly 'acquired by means of some kind of rational cognitive process.' The problem with Fodor's proposal, however, is that he takes 'acquired by means of a rational cognitive process' to be equivalent to 'learned by means of hypothesis testing.'

We have seen that hypothesis testing is not a very promising explanation of concept acquisition. And now it would turn out that if we reject it we automatically become radical concept nativists!

Samuels' proposal can perhaps be seen as a modification of Fodor's in that Samuels does not think that hypothesis testing is the only kind of psychological process that could be used in the explanation of concept acquisition. Still, even if we accept Samuels' primitivist definition of 'innate' as described above, it is not obvious that

---

173 Hypothesis testing is really the only rational cognitive or rational psychological process that Fodor considers as a possible way in which we acquire concepts. In various places he also talks about association (cf. for instance Fodor, 1983, pp. 29–31), but does not consider it as a possible process of concept acquisition. (Fodor talks about associationism when he discusses various theories of thinking. In that context, he rejects associationism as inadequate; he believes that "associative theories of thinking can't be right, since association doesn't preserve either sense or reference (to say nothing of truth) and thinking typically preserves all three," Fodor, 2008, p. 63.)

297
concepts are innate in the sense that they are 'already there, waiting to be triggered' as Fodor would like them to be. Rather, what this definition of innateness seems to suggest is only that we don't know how concepts are acquired. This is what Cowie calls 'the impossibility arguments' (Cowie, 1999, pp. 49–68). Arguments of this sort usually refer to the poverty of stimulus phenomenon as it was described above. Poverty of stimulus certainly does suggest something innate. It seems, however, that it could be explained as well by pointing to an innate mechanism which only requires partial and seemingly insufficient data in order to trigger a given feature.

Moreover, Samuels' proposal faces its own problem. If something is 'innate' just because it is 'not acquired by any psychological process,' then it seems that things like AIDS should count as innate (after all, AIDS is not acquired by any psychological process). Samuels tries to block this possibility by suggesting that his definition only applies to cognitive structures. Such a move makes sense, he believes, because there are good reasons to think that "the notion of innateness used by cognitive scientists should be understood (...) in terms of concepts that derive from psychology or even from cognitive science itself" (Samuels, 2004, p. 139). However credible these reasons might be, still, given that 'innate' is a term borrowed from biology, it seems that it better not be limited just to cognitive structures.

Perhaps, then, we should focus on a more positive account of innateness, that is, try to figure out how an innate trait is acquired, and not just how it is not. Notice that Fodor never provides any substantial explanation of what triggering is supposed to mean: he only says that whatever is not acquired by means of hypothesis testing, or
present from birth, is triggered. How could triggering itself be explained? One way to answer this question is by means of the concept of a catalyst (cf. Stich, 1975, p. 14). A catalyst is a necessary element for a chemical reaction to occur, but it is not a part of the end product of the reaction. Similarly we could say that when an innate trait is acquired, what triggers it is necessary for the acquisition of the trait, but it is not what supplies the content of the end product. That content, therefore, must come from the organism, and that's why it is called innate. (In negative terms, triggering is the opposite of teaching and of other kinds of acquisition, psychological or non-psychological, in which the content at the output (the content of what is acquired) is provided by (or similar to) the content already present at the input (the content of the stimulus). Compare, for instance, innate genetic diseases, such as autism, and STIs. In the first case, the disease is innate, even though it still requires some kind of environmental stimulus to be triggered (the presence of mercury in a vaccine or in mother's diet). The stimulus does not bring with itself, but rather triggers, the disease. The situation is different with STIs which are passed from another individual already having the disease.

In another, but quite similar way, triggering can be explained as an input-output kind of process in which the input and the output differ in content. The usual view on the difference between input and output in the triggering relation is that the input has a more limited informational content than the output. And so, we have, for instance, Chomsky who interprets the mind as an input-output device with sensory data as the input, and knowledge, beliefs or concepts as the output (Chomsky, 1975b, p. 122). He describes the process of belief acquisition as an input-output process such that
Q129. if the beliefs [i.e., the output] that result from a particular pattern of sensory experience [the input] are richer or contain more information than the experience, then this added information must be the mind's contribution [it must be innate]. (Stich, 1975, p. 15)

This interpretation of Chomsky's view seems to work well as an explanation for the phenomenon of language acquisition in general as can be seen, again, in the poverty of stimulus arguments for innateness. Little children, it is said, have a very limited experience with language. In particular, the grammar they learn is underdetermined by the instances of grammar they are presented with. And yet children become sophisticated language users very quickly. And so, it seems that experience alone cannot account for language acquisition. As Chomsky explains, "the data available as input is only a minute sample of the linguistic material that has been thoroughly mastered" (Chomsky, 1975b, p. 123). This is because the limited experiential stimuli only provide an occasion or a cause for the innate knowledge of universal grammar to manifest itself. (And so, going back to the input-output device analogy, we can conclude that "the basic properties of the output are a consequence of the design of the device," Chomsky, 1975b, p. 127).

However, there are certain problems here. First of all, the innateness that Chomsky has in mind seems to be rather architectural: it is the mind, or the language device itself that needs to have an innate structure and functioning, there is no need here for the presence of any innate representational content. The language organ does not need to include any innate knowledge of universal grammar, which, together with knowledge of observed language use and general rules of inference, would lead to
knowledge of a particular grammar. Instead, it may consist of a set of specialized
inference rules, which lead from the limited observational data to knowledge of a
particular grammar.

Chomsky's explanation of language acquisition does not explain the kind of
'triggering' that Fodor has in mind. The main idea behind Fodor's 'triggering' is that
something innate, some innate representational content that will show up at the output is
innate because gets triggered by some environmental stimuli. And this is not what
Chomsky is talking about.

Secondly, the very idea of 'the amount of informational content' is rather
problematic. It is not clear how we are supposed to measure and compare different
 informational contents. If we want to figure out whether, for instance, our sensory
concepts are innate, how would we measure the informational content of the concept
RED and compare it to the informational content of the experience of red objects? It
seems that on the discussed account there could be no such thing as innate sensory
concepts, because it is possible that our experiences contain much more information
than the concepts we acquire from them. (Perhaps the most abstract concepts, such as
GOD, the UNIVERSE, or INFINITY would have a better chance to count as innate.)
What's even more counterintuitive, on this scenario empirical generalizations would
also count as innate because "the evidence a person has for an empirical belief rarely
entails the belief" (Stich, 1975, p. 16).

It is important to notice that the latter two accounts, according to which innate
traits are triggered, suggest that different traits can be innate to different degrees: the
greater the disparity between the informational content of the input and of the output, the more innate a given trait. I'll put this issue aside for the moment. The conclusion that I want to make now is that whichever version of the triggering account of innateness we accept, the troubles with the definition of the term 'innate' are not over. 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 1981b). We don't want to say that the knowledge of Latin is innate. But it would have to be innate on all three explanations of triggering presented above. Samuels would have to say that the person's knowledge of Latin is innate in virtue of being a cognitive structure whose acquisition has no psychological explanation. For Fodor, the person's knowledge of Latin is not acquired by means of hypothesis 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. Finally, Chomsky would also have to say that the knowledge of Latin is innate: the output of the process has a very rich informational content and the input probably no informational content at all. This means that the whole informational content must have been provided by the person herself.

(iv) Innate III: The Normalcy condition

Samuels' solution to the Latin pill problem (Samuels, 2002, p. 257) is to come up with an additional constraint (an additional necessary condition) on what it means to be innate.
And so, a cognitive structure is innate, on this modified version of Samuels' primitivism, if it satisfies the normalcy constraint according to which

Q130. A (token) cognitive structure S possessed by an organism O is innate only if O would acquire S (...) in the normal course of events. (Samuels, 2002, p. 257)

And so, we would say that our capacity to walk is innate because it successfully occurs whenever a baby is raised in normal circumstances. On the other hand, our knowledge of calculus is not innate, because it won't occur unless special circumstances (learning in a classroom) take place, and it occurs by means of psychological processes of teaching, learning, reading, studying, etc. The knowledge of Latin from the example above is not innate on this modified version of Samuels' proposal because it is acquired in a very unusual way.

The concept of 'normalcy,' however, brings 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. Also, it would seem that 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. (A similar problem will threaten the definition of innateness in terms of genetic determination. I'll get back to it below, 5.2.3.)

In any case, 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, reference to some kind of minimal understanding of 'normal circumstances' will be needed for any plausible definition of innateness. This is because for any living organism there are some minimal requirements that must be met so that it can function properly — without them, no traits, whether innate or not innate, will be acquired.

(v) Various meanings of the term 'innate' in biological sciences

Samuels' 'normalcy condition' takes us back to the scientific origins of the term 'innate,' that is, to biology. In order to figure out what counts as a normal course of events or normal circumstances for an organism O we have to be able to determine what kind of an organism O is, or what species it belongs to. This means that we'd end up talking of 'normal circumstances for normally developing members of species S,' or that we'd consider those traits to be innate which manifest themselves in all normally developing members of a population P. This, however, is only one among many different senses in which the term 'innate' is sometimes used in biological sciences. Consider the following list of the meanings of 'innate':
• 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 that 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).

The problem with these various characteristics of innate traits is that there is no agreement on how exactly to define them (e.g., how exactly we should define a trait which is genetically determined), and also that they 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).

I am not going to try to determine whether the situation is so hopeless that we should follow Griffiths' suggestion and stop using the term 'innate' for good (cf. Griffiths, 2002, p. 82). Instead, I'm going to use some of the explanations of innateness that have been used in biology and try to come up with an account that will be useful for the assessment of the plausibility of Fodor's mad-dog concept nativism.
5.2 Two questions: How and why a trait is acquired

My suggestion in what follows is that it will be helpful to distinguish two questions that need to be answered when trying to figure out whether a trait is innate. First of all, we need to explain how an organism ends up having the trait in question—how was it acquired? Secondly, an explanation is needed of why the organism has that trait in the first place. The answers to these questions will indicate certain desiderata that any plausible explanation of the term 'innate' will have to satisfy.

5.2.1 First Question: How are innate traits acquired?

The answer to the first question will have several parts. First of all, we will 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 didn't have it, and then 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:

    Innate Desideratum 1:
    An innate trait may not 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 couldn't 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 doesn't count as innate.)

Innate Desideratum 2
An innate trait is not learned.

Third, we will need to figure out what factors are necessary for the trait to be acquired. Since all features in our phenotype depend basically on two things, our genetic makeup and, 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 both genetically determined, and require certain environmental factors. And so, we have two more desiderata (Desideratum 4 will be explained in more detail below):

Innate Desideratum 3:
The possession/acquisition of an innate trait does require and depend on environmental factors.

Innate Desideratum 4:
An innate trait is a trait that is somehow transmitted genetically. It is somehow coded in the genes and in normal circumstances all normally developing members of a given species are going to possess it without the need for any special circumstances (it is not unconditionally guaranteed that the organism will actually acquire a genetically determined trait; serious damage may prevent acquisition of the trait, etc.).
All the four desiderata listed above are necessary but not sufficient for innateness.

Whether a trait is innate or not will be decided once we answer the second question.

5.2.2 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 the main question that we discuss is how concepts are acquired by the human thinkers, what we really need is a definition of a trait that is innate in a species. (There may be traits produced by random mutations in an individual organism, and then inherited by its offspring, but such traits are not the topic of the present considerations.) For this reason it makes sense that in order to answer the question why an organism of a given species has an innate trait that it does, we refer, in the end, to evolution. And so, we'll say that

Innate Desideratum 5:
An innate trait is a trait that is produced by natural selection; it was selected because of some survival advantage for the species, e.g., because it was useful for enhancing control over behavioral responses to environmental stimuli.

Let me now try to provide a more thorough explanation of the two crucial points of my proposal to the effect that innate traits are those traits that are genetically determined and that were produced by natural selection (that are evolutionary adaptations).
5.2.3 Innate IV: Genetically determined

Genetic determination of a trait can be explained in various ways (cf. Samuels, 2004, pp. 137–8 and Symons, 1992, pp. 140–2). We may say, for instance, that a trait is genetically determined if it is caused by the genes. There are two possibilities here. One is to define an innate trait as a trait that is caused entirely by genetic factors. This, however, is not very fortunate: it would imply that there are no innate traits at all, because all traits require interaction between the genes and the environment (cf. Samuels, 2004, p. 138).

Alternatively, if just having a genetic cause is sufficient for something to be innate, then it would turn out that all traits of an organism are innate because in some sense all traits that an organism possesses have genetic causes (cf. MacLaurin, 2002, p. 106; for an example, think of rollerblading, which is certainly not an innate capacity, but does require some genetic basis). Similarly, if we try to define an innate trait as a trait which causally co-varies with some genetic pattern, this again won't work because there are no traits that co-vary with nothing but some genetic patterns, but rather "all traits causally co-vary with both genetic and environmental features" (Samuels, 2004, p. 138).

Another explanation of genetic determination to consider seems similar to Samuels' normalcy condition as described above. On this proposal, "a phenotypic trait is innate for a given genotype if and only if that phenotype will emerge in all of a range of developmental environments" (Sober, 1999, p. 795). In other words, a trait is innate 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 changes (with the assumption that the changes can only be such that they would still allow the organism to survive and thrive). 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 traits that are developed by normal members of a given genotype would count as innate.

The problem with this third explanation of genetic determination is, as it was the case with the normalcy condition discussed above (5.1.4), that it would also lead us to identify too many traits as innate. It seems plausible that all normal adult human beings in a variety of environments have the belief 'Water is wet.' So, this belief is invariant. But we would not want to say that it is innate. And so, it is obvious that invariance on its own is not sufficient for innateness. Samuels offers a diagnosis of the problem when he explains:

Q131. 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)

I agree with Samuels' diagnosis. I believe, however, that there is a better answer to the problem than the psychological primitivism that he offers (cf. 5.1.3). In fact, Samuels' primitivism does not solve the problem but only pushes it back. If we apply the invariance account of genetic determination to mental representations, we'd say that those concepts (or beliefs) are innate that develop 'in all of a range of developmental environments.' To this Samuels would respond: this does not work, because it only
shows that there is a mapping relation between the human genome and the phenomenon of concept (or belief) acquisition. What we need to do in addition, Samuels would claim, is to explain that concepts or beliefs are innate only in cases where such a mapping is a result of the fact that we all share some non-psychological mechanisms by means of which we acquire them. As we have seen above, however, there is a problem with this proposal. We do not know how to identify just those non-psychological processes that indeed allow us to acquire innate mental representations. Also, if we accept Samuels' suggestion, we would end up with non-psychological mechanisms which themselves are invariant phenotypic traits that occur in organisms with a certain genotype. Again, we would have a mapping relation between genotypes and phenotypes, without an explanation of why such a mapping takes place. So, a different approach is needed.

What we need to explain is indeed why/in what sense a given trait is invariant. There will be two possible answers to this question. In some cases, a trait is invariant because of environmental similarities (an example of such a trait is the belief that all humans share that water is wet). In other cases we can talk of some traits being innate: those are the traits that are invariant because they are evolutionary adaptations.

5.2.4 Innate V: Evolutionary adaptations

The proposal according to which those traits are innate that are genetically determined in the sense of being evolutionary adaptations seems to have an important advantage. It
allows us to really answer the question that we ask when we try to figure out whether some trait is innate. We are not just interested in figuring out how we end up having a trait in question. We are perhaps even more concerned with the question of why we have it at all.

To claim that a trait is an adaptation means

Q132. 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)

Now, we need to be clear on what kind of reference to evolution is needed for our purposes. On my account, innate traits are not those features which have arisen "from a random evolutionary force like a big mutation, statistical drift, or the fortuitous shape of the nooks and crannies between [various] organs" (Pinker, 1997, p. 36). Rather, they were produced by natural selection. Natural selection is not random. The reason why certain traits are selected is that they provided an important survival advantage to members of a given species. Natural selection

Q133. cares only about the long-term fate of entities that replicate; that is, entities that retain a stable identity across many generations of copying. It predicts only that replicators whose effects tend to enhance the probability of their own replication come to predominate (Pinker, 1997, p. 43)

Natural selection, as Pinker explains, is "the only evolutionary force that acts like an engineer 'designing' organs that accomplish (...) adaptive outcomes"; it is "the only non-
miraculous process we know of that can manufacture well functioning [biological] machines" (Pinker, 1997, pp. 36–7).

And so, on my proposal an innate trait is defined as a trait that is genetically determined in the sense that it is an evolutionary adaptation; it was produced by natural selection, and fixed in a given species because of its survival value (cf. Symons, 1992, p. 141).

Here is how the connection between natural selection and genes is explained (cf. Williams, 1996, and, Futuyma, 2005, esp. chapters 11 and 12). From one generation to the next there are changes in the genetic material of a given population of organisms. Genes are passed from one generation to another. Mostly by means of random genetic mutations a variation of inherited traits is produced. Some individuals possess advantageous or adaptive traits, and as a result are more likely to reproduce. In the next generation, there will be more individuals possessing the adaptive traits in question. This whole process is what is called natural selection. (Another process that influences what traits become more common in a population is called genetic drift; I ignore it in this dissertation for the reasons of brevity.)

It is important at this point to make sure that we distinguish 'evolutionary adaptations' from 'adaptive traits.' Adaptations are those traits that evolved by natural selection; they occur through a combination of successive, small, random changes in traits, and become fixed in a population by natural 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 of the species (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 in time. Language, for instance, can be considered an innate feature of human beings for the last several thousands of years. But there has been a certain time when this trait appeared as a result of just a random mutation. After that, at a certain time it turned out to be an adaptive trait—creatures endowed with language were more likely to reproduce. Now, as we said, language can be considered an evolutionary adaptation.174

It seems that what we’ll get in the end is three classes of traits that could be considered as genetically determined.175 First of all, we’ll have (a) innate traits, that is, those traits that are inherited and that are evolutionary adaptations. Among the non-innate traits we’ll distinguish (b) those traits that are inherited, but that were not selected for their survival value (e.g., because they didn't have any), and (c) the traits that still

---

174 This definition of innateness also indicates, as it was the case with the definition in terms of triggering, that different traits may be innate to a different degree. Perhaps the way it works is that 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 than other traits that the old parts of the brain are responsible for.

175 According to MacLaurin (MacLaurin, 2002, p. 110), we should not limit innate traits to those that are genetically determined evolutionary adaptations. MacLaurin rejects the account according to which innate traits would be defined as ‘products of genetic information that are fixed (in the sense of being maintained by various mechanisms) in some biological populations.' MacLaurin agrees that the important function of our genes is that they code for phenotypic traits (what genes you have will influence what observable features you’ll have) and that they program development (your genetic makeup determine the speed and other factors of how you develop). He emphasizes, however (and this is the common view among contemporary developmental systems theorists), that genes are not the only factors that play these roles. There are various non-genetic developmental resources (nutritional, cytological, embryological and behavioral) that influence the organism’s development and the phenotype it’s going to have. (Different developmental resources provide information on what your phenotype is going to be. There are some elements of this phenotype that are fixed because of their survival value even though their causes are mostly non-genetic.) So, according to MacLaurin it would be better to define innate traits as those traits that are products not just of genetic information, but rather of information provided by various developmental resources that "are maintained in biological populations by a variety of mechanisms" (MacLaurin, 2002, p. 110). I don't want to go this far into details. So I'll just stick with the idea that a trait is innate if it is a genetically determined evolutionary adaptation.
require some genetic foundation (as all traits do), but that are acquired mostly as a result of environmental stimulation (they would not appear without some specific kind of stimulation coming from the external world).

Here is how various traits would be classified on my proposal:

Innate

(a) All those traits that are both inherited and evolutionary adaptations; that are fixed in a given population; and that were chosen by natural selection because of their survival value. It is possible that a trait that is an evolutionary adaptation may not at present have the same (or any) survival value for the species (it may not be an adaptive trait). The following traits would count as innate on this definition:

- Chomsky's language organ
- Modularity of perception/language
- Functioning of the heart/circulatory system
- Possession of five fingers by humans
- Walking
- Capacity of infants to quickly acquire a theory of mind
- Human taste for sugar
- Various atavistic traits

Non-Innate

(b) Among traits that would count as non-innate, we'd include those traits that are still under a quite close genetic control, that were inherited but not fixed in a given population because they didn't have any serious survival value.

- Eye color
- The exact length of a person's fingers
- Innate diseases (although, see below, note iv)

(c) Random/More Dependent on Environment

- Capacity to acquire the belief that water is wet
- Capacity to learn French as a second language
Taste for green tea
Capacity to learn how to rollerblade (how good one could become is determined by the genes to a large degree; it would probably belong to group (b))

Here are some additional implications that follow from the definition of the term 'innate' that I offer.

(i) Currently adaptive traits may count as either innate or non-innate.
(ii) Whether a trait is an adaptation is an empirical matter. For this reason, we may expect not to know sometimes whether, and if yes, to what degree, a trait is innate (a trait would be considered more innate if it's been an adaptation for a longer period of time, it is 'more fixed' in a given population).
(iii) Some non-innate traits that are at present advantageous for the survival of the species, and that are also heritable, may become innate (of course this would most likely take thousands of years).
(iv) If a trait that is innate, that is, a trait which is an evolutionary adaptation, stops being an adaptive trait, it may still remain fixed in a given population, even though its survival value diminishes or disappears. Such a trait, even though it is not at present an adaptive trait, still counts as innate. For instance there is what we'd consider an innate propensity for elevated levels of glucose, glycerol and other sugar derivatives in some human populations of northern Europe (especially in Scandinavia). This feature is responsible for a higher prevalence of Type 1 diabetes in those populations, and so (now) it is disadvantageous. Nonetheless, scientists hypothesize that factors predisposing to higher concentrations of these substances may have been selected for as adaptive measures in exceedingly cold climates (elevated levels of sugar derivatives depress the freezing point of body fluids and prevent the formation of ice crystals in cells through supercooling, and so, they constitute a cryoprotective adaptation, an

adaptation that protects from very cold temperatures). In times when life expectancy was short, factors predisposing to Type 1 diabetes provided a survival advantage. In more modern times, in a significantly warmer climate where the cold is not as threatening, and as life expectancy has increased, this innate condition (at least assuming we're not dealing with the late onset of diabetes) has deleterious consequences, outweighing its protective value.

(v) For a trait to stop being innate it really means for it to disappear. For instance, assuming that those evolutionary theories are right that hold that early humans were ape-like to a great extent, we'd conclude that humans used to be much more skillful at climbing trees, that is, that humans, as well as other species of primates, had an adaptation for climbing trees. Such skills (together with various physiological details that facilitated them, e.g., the special structure of the ankles, the possession of tail, etc.) used to be innate traits of humans (in the same way in which they still are innate for monkeys). But then humans stopped to climb trees. Tree climbing skills stopped being crucial for the survival of the human species, and they greatly diminished (we don't have a tail anymore, etc.).

5.2.5 Fodor and Aquinas

Most importantly,

(vi) The general definition of innateness that I propose is compatible with both Aquinas's and Fodor's views. Both philosophers would agree (assuming the relevant scientific data was available also to Aquinas) with the classification of the traits listed in group (a) above as innate. Also, both philosophers could accept the five desiderata for
innateness described above. (It will be a separate question whether they would or
should conclude that concepts as thought-parts are innate. Cf. Conclusion.)

Desideratum 1
'Innate' does not mean 'present at birth.'

This is nowadays a commonly accepted view. Fodor would not have any reason to reject it.
Aquinas, as it seems, could accept it as well. On his view, what's innate for a thing is its
'whatness,' or its essence. It is innately determined what kind of thing an organism is supposed
to be, what features, assuming that all goes well, it has to develop as its proper accidents
(features that express its essence), and also what non-proper accidents it has the potential to
develop. All of this is 'written up' as a recipe in the organism's essence. This essence is far from
being realized at birth; rather, an organism will strive to realize its essence as it grows and
develops. And this means that on Aquinas's view, many of innate traits of an organism are not
yet present at birth (even though they are 'there' in potentiality).

Desideratum 2
Innate is opposed to learned.

This desideratum is just definitional, and so I'll assume it to be uncontroversial.

Desideratum 3
Innate traits may and usually do require sensory experience.

Fodor's account of innateness in terms of triggering (see above, 4.3) suggests that he'd agree
with this third desideratum: triggering is needed for acquisition of innate concepts, and
triggering involves sensory experience. Aquinas wouldn't have a problem here either. He holds
that sensory experience is required for all rational animals to reach their nature's perfection;
what this nature is, and what its perfection consists in is innate, but, for embodied beings, there
is no way to reach it without sensory experience.

Desideratum 4
'Innate' means 'genetically determined.'
This desideratum is explicit in Fodor's account of input systems (see above, 4.4.1). How the modules are structured and how they function has to be specified by our genetic programming. More generally, the fact that we all resemble each other is due to similarity of our genome.

Aquinas of course didn't know anything about the genes. So, he would not have a scientific explanation of how exactly, by what factors it happens that the structure of the mind is the way it is. He would certainly admit that God must have some means to ensure that members of the same species resemble each other, and that parents transmit various traits to their children. Also, recall that above (1.1.3) we suggested an interpretation of Aquinas according to which what Aquinas calls 'substantial form' could be understood as a non-physical counterpart of the genetic code. Substantial form of an organism would then be a pattern, a code in which all the necessary information about it is written. We could say that offspring resembles its parents in virtue of the information it has received that was written up in the parents' substantial forms.

Finally, Desideratum 5
An innate trait is an evolutionary adaptation.

One view that Fodor considers in his LOT2 is that the learning of stereotypes (or prototypes) is the first stage of concept acquisition. It is only after having acquired a given stereotype that a person is able to grasp a corresponding concept (Fodor, 2008, p. 162). The grasping of the concept occurs by means of some innate disposition. Fodor's idea that acquiring a prototype is needed as a prior stage to acquiring a concept is quite problematic (see above, 3.3). Still, the whole view that he considers can be interpreted as indicating a correct intuition to the effect that there really is no way to get rid of empiricism in (basic cases of) cognition. There is no way to arrive at a PROTOTYPE DOG (assuming that we exclude acquiring it by means of reference borrowing) without sensory experience; this is because we cognize properties of things by means of our sense organs. At the same time, it has to be by means of some innate disposition that perception of an object leads us to a given concept. This innate disposition may be an innate mechanism such that sensory experience of a certain object makes it happen that an atomic symbol of the brain code is recruited, a specific neural pattern is produced whose firing will constitute thinking a given concept. This innate disposition would be an evolutionary adaptation. (In Fodor's consideration of concept acquisition there is no suggestion that innate
traits are those that are evolutionary adaptations; in other contexts (cf. e.g., Fodor, 2001, p. 102), however, he does show his appreciation for a biological notion of innateness. In any case, there are no apparent reasons for him to reject the version of a biological definition of innateness that refers to evolutionary adaptations.)

Aquinas, as we said, didn't know anything about genes. In addition, he followed Aristotle in thinking that "nature is among the number of causes which act for the sake of something" (In II Phys 12, 250). He had a teleological view of nature, a view that is often considered to be inconsistent with evolution and natural selection, that is, with the two processes that don't really have any 'goal.' For Aquinas, nature itself is teleological, that is, nothing it produces is useless. All living things have a function, and a final cause. That is, each of them can be explained in terms of the goal for the sake of which it's been created. A thing's final cause is closely related to its formal cause, i.e., its essential nature. For living organisms their final cause is simply the fullest realization of the formal cause, that is, realization of their true nature. In addition, on Aquinas's views, the instincts to survive and to reproduce are indeed the most powerful and the most fundamental instincts for all living things. It seems plausible that if Aquinas knew about genetic mutations, about the resulting population variations with respect to various traits, about environmental changes and variations, and about natural selection, he could agree that this is what God has planned as a mechanism of introducing improvement to its creation, and of adapting it to the changes in environment.

What we need at this point is to see how the proposed definition of innateness could be applied in the case of representational nativism. We need an explanation of what it would mean for a representation to be innate, what it would mean to say that some kinds of...

---

Aristotle in his Physics considered a view that sounds very similar to natural selection (see Aristotle, Physics II, 2, 198b20-33: "Why then should it not be the same with the parts in nature, e.g. that our teeth should come up of necessity—the front teeth sharp, fitted for tearing; the molars broad and useful for grinding down the food—since they did not arise for this end, but it was merely a coincident result; and so with all other parts in which we suppose that there is purpose? Wherever then all the parts came about just what they would have been if they had come be for an end, such things survived, being organized spontaneously in a fitting way; whereas those which grew otherwise perished and continue to perish."), but didn't further develop it.
representations are innate in the sense of being genetically determined evolutionary adaptations.

5.3 Representational innateness

5.3.1 Invariance Tissue Nativism

Elman and his co-authors (Elman, 1996) consider one version of representational innateness, that I call Invariance Tissue Nativism (ITN). On this view, to say that there is innate knowledge is to claim that some representations are hard-wired into the brain. What this means is that various neurons are "born 'knowing' what kinds of representations they are destined to take on" (Elman, 1996, pp. 26–7), that is, they lack developmental plasticity. More generally, ITN 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 for vision. On the definition of innateness that I propose, ITN would also claim that this pre-specified knowledge that neurons possess is an evolutionary adaptation; it is genetically determined because of its survival value for the human species.
Elman and his co-authors believe that this version of RN should be rejected. Their evidence against ITN comes from various studies with vertebrate animals. In some of the studies, parts of the cortical tissue of an animal are transplanted from one area to another. In other studies, sensory input is redirected to an unexpected area of the cortex, e.g., retinal inputs are projected onto the auditory cortex. In the first case, it turns out that the transplanted tissue takes on the features and functions imposed on it by its new location. In the second case, inputs that are usually processed by the visual cortex are processed instead by the auditory cortex. Both experiments show that the cortical tissue is plastic: what representational structures are developed in a given part of the cortex is dependent, at least in part, on the stimuli incoming from the environment. This plasticity of cortical tissue supposedly proves that representational nativism does not hold.

5.3.2 Contra Elman's argument against ITN

i. Developmental plasticity does not preclude innateness

As we said above, to claim that innate traits are genetically determined in the sense that they lack developmental plasticity (or, in other words, that they are insensitive in their development to environmental factors; cf. Griffiths, 2002) is only one among different ways in which genetic determination can be explained. It is quite a different thing to define an innate trait as a trait that is an evolutionary adaptation, or as a trait that the organism is designed to acquire because of the survival value that it has for the species.
Lack of developmental plasticity with respect to a certain trait does not imply that the trait is an evolutionary adaptation fixed in a given population (cf. Griffiths, 2002). Rather, a trait can be an evolutionary adaptation while being developmentally plastic. For instance, the capacity to distinguish colors is fixed in the human population in the sense that in proper circumstances all normal human beings will possess it. It also makes sense to say that it is an evolutionary adaptation, and that it fixed because of some survival advantage. But it is also developmentally plastic, in the sense that if we replace a piece of the visual cortex with the piece of auditory cortex, the new tissue will take on the functions of the original: the (previously) auditory cortex will become responsible for the capacity to distinguish colors. Similarly, the functioning of the human heart, which is another evolutionary adaptation, is fixed in that all (normally developed) human beings have it; but it's also developmentally plastic to the point that the heart tissue can be completely replaced with an artificial heart. (What's innate here is really the function that the heart plays—it pumps blood. This function is plastic, that is, multiply realizable.) And so, a trait can be developmentally plastic but still innate in the sense of being an evolutionary adaptation.

ii. Straw-man Nativism

According to Samuels, the ITN that Elman and his co-authors argue against is in fact nothing but a straw-man nativism (Samuels, 1998, p. 552). It doesn't make sense, as Samuels points out, because it seems that those who would hold it would have to be
committed to the view that "innate representations are somehow more innate than skins, hearts, teeth and the like." Samuels continues:

Q134. It would presumably be singularly implausible to conclude that a certain group of cells C (in a normally developing organism) is not innately specified to become, say, teeth, on the grounds that if C were transplanted to a different location in the organism early in development, then it would develop into (say) part of the belly region. (Samuels, 1998, p. 558)

ITN is indeed not very plausible. For this reason it is so easily refuted by Elman's experiments. It seems, however, that it would not be so hard to remedy the problem. (Samuels notices it as well.)

iii. The Modified Invariance Tissue Nativism

In order to avoid the problems indicated by Elman's argument, it would be sufficient, it seems, to add some kind of 'normalcy condition' to the Invariance Version of Tissue Nativism. We could hold that it is innately specified that in case of normal development and normal environmental stimulation, certain parts of the brain (will) have certain specific functions, and that certain parts of the cortex are (or will be) responsible, for instance, for vision, but also that it is equally innately determined that if parts of visual cortex get transplanted to a different part of the brain, they may take upon them different functions.

This formulation of ITN will not be immediately refuted by Elman's kind of evidence. It will always be possible that the cortex develops abnormally, or that it is not normally connected to other parts of the brain, or that parts of it are transplanted into a
different place in the body. If this happens, we would not expect the cortex to behave normally.

iv. Fodor and Aquinas on ITN

It seems that both Fodor and Aquinas would accept the claims of the modified ITN. Fodor would have no reason to deny that it’s genetically determined and an evolutionary adaptation that at a certain point in the brain development some cells are specified to become, say, visual cortex, and that in normal circumstances, if nothing goes wrong, they will become exactly that. As concerns Aquinas, he believes in a teleological nature of everything that exists. He’d agree, therefore, that all parts of the human body, also at the cellular level, have some pre-specified functions. If nothing goes wrong, all parts of the human body will end up realizing their proper functions.

In any case, the modified ITN on its own does not imply anything about concepts being or not being innate. It may be that there is a specific part of the brain which is responsible, for instance, for chess playing skills—if this part of the brain is missing, chess playing becomes impossible. But this would not be enough to imply that concepts needed for chess playing are innate. The fact that some pieces of the brain are necessary in order for the subject to ever develop a specific capacity does not mean that the capacity is innate. Also, playing chess is clearly a learned capacity.
5.3.3 Tissue vs. Organism Representational Nativism

Samuels points to another problem with, this time, both formulations of invariance tissue nativism discussed above. Tissue nativism, Samuels explains, constitutes only a subcategory within representational nativism. And arguments that work against tissue nativism would not necessarily harm the other kind of representational nativism, that is, organism nativism (Samuels, 1998, pp. 559f). It is the latter, according to Samuels, that nativists about mental representations really have in mind.

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

Q135. 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)

Innate beliefs (or concepts), according to organism nativism, are in some way genetically programmed to arise in our mind; i.e., innate representations are the phenotypes that all humans (at least all those with similar enough genotype, and those who develop in normal conditions) have in common. In addition, on the definition of innateness that I propose, organism nativism would also claim that it is an evolutionary adaptation that all humans will acquire certain mental representations.

According to Samuels, even if Elman is right (that is, if we assume that nativism is to be defined in terms of lack of developmental plasticity, and that the experiments he refers to indeed support the claim that the cortex is highly plastic), his argument does
not work against "contemporary theorists who defend nativism about representations," Fodor included. Those theorists, Samuels holds, 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 processes) can be realized by different kinds of neural entity (state or process)" (Samuels, 1998, p. 560). And tissue-RN, according to Samuels, wants to deny multiple realizability.

The first problem with organism nativism, however, is that, as it was the case with other theories referring to the normalcy condition, it would 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 don't 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 that (2) what kind of concepts it will acquire will be 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 sure, Fodor is a
nativist. And so is Aquinas. Organism nativism with respect to various mental representations does make sense. 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 this requires innate propositional knowledge, or innate thought-parts.

5.3.4 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'll find it—contrary to Samuels' claim—on the side of tissue nativism. We should notice, however, that in his criticism of Elman, Samuels switches back and forth, without noticing it, between two distinct versions of tissue nativism. Both of these versions are suggested, although not explicitly, by Elman's definition of innateness.

Elman's definition of representational nativism included two claims:

Elman 1: Various neurons are 'born knowing' what kinds of representations they are destined to take on.

and
Elman 2: Some representations are hard-wired into the brain; that is, they are encoded in the brain as particular patterns of synaptic connectivity within a specific neural system. (Elman, 1996, pp. 26–7)

These two claims point to two distinct kinds of tissue nativism. The first claim expresses what I called Invariance Tissue Nativism (ITN). It can be explained, as we have seen, as a more plausible and a more general claim to the effect 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 (if something goes wrong, the pieces of tissue are then genetically determined to take on another function, etc.). The second claim expresses what I'll call Neural Nativism (NN).

Neural Nativism (NN) 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.

178 According to NN, those pre-specified neural structures are genetically determined to represent specific objects: 179 it has been inherited by the individual and evolved in the

178 Here is how a similar view could be formulated in connectionist terms: "In a connectionist network, representations are patterns of activations across a pool of neuron-like processing units. The form of these activation patterns is determined by the nature of the connections between the units. Thus, innate representational knowledge—by which we mean the potential to produce representations of specific sorts—would take the form of pre-specified weights on the inter-unit connections" (Elman, 1996, p. 25).

179 Samuels seems to be referring to this kind of tissue nativism when he considers the view according to which "specific pieces of cortical tissue are innately specified to encode certain representations" (Samuels, 1998, p. 559), and also says that "It might be innately specified that a specific cluster of neurons in the prefrontal lobe encode the representation [DOG]" (Samuels, 1998, p. 559), or that "it is innately specified that some cortical structure realizes a given mental representation R" (Samuels, 1998, p. 561). But then Samuels
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.

Explained this way, Neural Nativism would support Fodor's mad-dog nativism. Fodor's nativism would be the view according to which 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.

This seems to imply that the way to interpret Fodor's view is as

Q136. 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. 561–2)

On Fodor's mad-dog nativism, all possible types of primitive mental representations (cf. above, 4.1.2 and 4.2) would be in advance assigned to specific types of neural patterns in specific locations in the brain. All those types of neural patterns would be hard-wired;

---

proceeds as if there was no difference between what it means to genetically encode the tissue, and what it means to genetically encode a representation. He does not distinguish the view according to which specific pieces of tissue are genetically determined to take on certain functions (ITN) from the view according to which particular neural patterns that get activated in specific sets of neurons in a specific location of the brain are genetically determined to represent specific objects (NN). (Notice that one could hold to Neural Nativism in either case, whether developmental plasticity is or is not true.)
they would be inherited by the individual and fixed in the human species because of the adaptive value of that feature.

At this point there are two questions that need to be answered. First of all, it is necessary to figure out to what extent Neural Nativism is plausible from the scientific point of view. Secondly, and more importantly, we need to consider whether it makes sense for Fodor to hold such a view given his account of the architecture of the mind and his account of concepts.

5.3.5 What would science say?

It is pretty much an accepted view these days that "everything humans know (and do) is served by and represented in the human brain" (Elman, 1996, p. 241), and that different things that we know are stored in the brain in non identical ways (one day, with the help of some future brain-imaging machines, we may, in principle, be able to see how different representations are stored in the brain).

As we said, Neural Nativism that 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 will 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 will represent any given
object is genetically determined and fixed in advance by natural selection, independently of experience.

In order to explain the process by means of which our genes would code for such innate primitive mental representations we could perhaps refer to what's called the Hebbian theory of cell interaction. On this theory,

Q137. 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

Q138. 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.")

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 distinguish different concepts at the neural level, we would need 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 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.

How plausible is such a view? It seems that at this point of its development science is not able to determine whether the view that Neural Nativism promotes is true. What scientists can determine these days is which parts of the brain are (the most) active during various cognitive tasks (cf. Elman, 1996, pp. 3–5). (This way, that is, because they are able to determine what part of the brain will take upon itself what function, assuming that everything goes well with the organism's development, they should, at least in principle, be able to provide evidence in support of various claims expressing the Modified ITN.) They have no way to say, however, what exactly happens in the brain when I think 'CAT' (cf. Elman, 1996, p. 4, and Merritt, 2008, p. 184). And so, they have no way to tell us whether any kind of concept nativism, whether extreme or partial, is the case.

What science does seem to suggest is that (1) specified localization is not sufficient for innateness, and that (2) neither specified localization nor identity of neural patterns across individuals are necessary for innateness. So, first of all it turns out that 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. It turns out, for instance, that there is a specific region of
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 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.

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,

Q139. Brain 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)

So, there is no need to claim 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.

Most importantly, however, 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. 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
function of the heart, of the circulatory system, or of visual edge detectors. This would
clearly be advantageous that the brain does 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 visual cortex detect edges. Even here, however, the brain must acquire these patterns at a certain point in its development. In addition, it turns out that

Q140. 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 we should say that what's innate in these cases is that some neural patterns get hard-wired very early in the development of the organism, and that they have a high level of priority in the allocation of neural resources. The same does not seem to make sense as concerns mental representations.

(i) Neural Non-Nativism is an Evolutionary Adaptation

With respect to conceptual mental representations, 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, can 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 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). This flexibility allows us to learn a basically infinite number of different concepts; also, because of it we are able to learn throughout

\footnote{What I am concerned with is really only a subset of mental representations: conceptual representations (Fodor would call them computational mental representations). I do not intend to take a stand concerning the perceptual representations. It may well be that the basic representations involved in perception, such as edge and motion detectors, are evolutionary adaptations.}
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. It seems enough to claim that 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 the neurons, etc. There is no need for the genes to code for a specific neural pattern for every possible concept.181

(ii) If concepts are not innate, how do we acquire them?

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: the set of patterns that was chosen by natural selection because of its survival value for the species. Acquiring a non-innate trait, therefore, involves acquiring such patterns of synaptic connections which are not evolutionary adaptations. When I learn how to rollerblade, new patterns of synaptic connections are produced in my brain: the exact character of these patterns depends on various circumstances; but in any case, it was not pre-determined by natural selection to take care of the human capacity to rollerblade.

---

181 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."
Perhaps it's similar with concept acquisition. We may still not know how exactly we acquire concepts. And we may still not know how we get to concepts from sensation, or from percepts. But we may say that concepts themselves are not innate because when we acquire a new concept, we acquire in our brains a certain pattern of synaptic connections that is not an evolutionary adaptation. The capacity to acquire concepts is an evolutionary adaptation; but how particular concepts are realized 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 environment. 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.

To sum up, Samuels criticizes Elman for reducing representational innateness to tissue nativism (or, in my terminology, to the Invariance version of tissue nativism). This kind of nativism is, as Samuels suggested, a straw-man nativism, but at the same time, it is easy to correct. The Modified Version of Tissue Nativism is, however, not that interesting, and does not suggest concept nativism anyway. Samuels seems to think that it is organism nativism that Fodor (and other philosophers) usually have in mind. In my opinion, however, organism nativism should rather count as architectural. On the other hand, if Neural Nativism was true, it would confirm Fodor's mad-dog nativism. But
given what science tells us, it's not very plausible that Neural Nativism is true. On the contrary, it seems that we should conclude that mad-dog concept nativism should be rejected because it would be disadvantageous for the organism. At the same time, it would be not nearly as implausible to claim that there is a small group of basic concepts, such as, for instance, MOTHER, FOOD, or DANGER, which are innate.
VI. CONCLUSION

SHOULD FODOR LISTEN TO AQUINAS? THE OUTCOME OF THE DEBATE

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 particular concepts as thought-parts to be innate. What we need to figure out now is whether Neural Nativism makes sense for Fodor given his account of the architecture of the mind and his account of concepts.

i. Implications of Fodor's Account of the Architecture of the Mind

In Fodor's account of the architecture of the mind, transducers are temporally first in the process of cognition of things external to the mind. They are connectors between cognizing organisms and the external world. There is no cognition of the external world possible without transducers, and similarly no corresponding concepts without them. So obviously, sensory experience, external stimuli, and properly functioning transducers are necessary for us to acquire any sensory concepts at all. These are architectural constraints on sensation. The fact that transducers require external stimulation does not
yet imply that (sensory) concepts are not innate. It may well be the case, and this is what Fodor thinks is the case, that stimuli received by transducers only trigger concepts that are 'already there.' On the other hand, Fodor's account of transducers does not indicate any need for innate concepts. Concepts themselves don't yet enter the stage at this point in the process of cognition. And 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, it appears that 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 modularity thesis. Fodor defines a module most generally as "a piece of hardware (...) whose structure somehow imposes limitations upon its capacities" (Fodor, 1983, p. 8). Each module, he says, is associated with a fixed neural architecture (Fodor, 1983, pp. 98–99).

Q141. The intimate association of modular systems with neural hardwiring is pretty much what you would expect given the assumption that the key to modularity is informational encapsulation. Presumably, hardwired connections indicate privileged paths of informational

---

382 On Fodor's description of transducers, it is not clear whether they are needed for acquisition of concepts of our inner states. It seems that they could as well be needed. All changes in both the external and the internal environment are the source of stimuli received by the transducers. Some stimuli carry information from the external environment (e.g., light or sound waves); other stimuli carry information about the internal environment (e.g., different kinds of information from interoceptive or proprioceptive senses). Any information that is relevant for what the organism will be transduced, translated in the format that the cognitive faculties will understand. As we have seen, Aquinas would most likely share this view. For Aquinas, any kind of knowledge, and acquisition of any concept, is impossible without some kind of contribution of the external senses. In any case, I'll leave this issue aside, and focus on such concepts as DOG, MOTHER or DOORKNOB.
access; the effect of hardwiring is thus to facilitate the flow of information from one neural structure to another. (Fodor, 1983, p. 98)

This is, again, architectural nativism. But then, Fodor also talks about modules having 'knowledge' and 'beliefs,' and 'making inferences.' He claims that "much of the information at the disposal of [input] systems is innately specified" (Fodor, 1983, pp. 100–1), and he says explicitly:

> Q142. 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)

Each module, according to Fodor, "comes with a database which is, in effect, what it innately believes about its proprietary computational domain" (Fodor, 2000, p. 91).

Here is how Fodor explains the innateness of the input systems' databases. The reason why the modules are useful in contributing to a creature's fitness is, Fodor admits, that 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 among the beliefs that the mind holds some 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 explains 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). Currently innate beliefs in the modules' databases, according to Fodor, used to be 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 trait itself got fixed as an evolutionary adaptation.\textsuperscript{183}

Still, there seems to be no support here for the kind of nativism which would imply that concepts as thought-parts are innate. Nothing in the evidence Fodor cites supports his mad-dog representational nativism over an architectural nativism according to which the modules 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 implying organism \textit{nativism} which is more architectural than representational. It is the kind of the 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.

Another reason why it seems that Fodor’s account of input systems does not imply representational nativism comes from his account of reflexes. For Fodor, reflexes are a good example of an innate faculty. He explains innateness of reflexes in terms of how they know \textit{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 \textit{concepts} are the innate element

\textsuperscript{183} 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).
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 addition, consider Fodor's critique of Chomsky in which Fodor himself suggests that it's only the architectural kind of nativism that Chomsky's views seem to imply. Nothing suggests the need for innate concepts or innate representations.

ii. Fodor against Chomsky

In his *Modularity of Mind* (1983), Fodor considers himself to be a faculty psychologist of the kind that holds that what's innate in human cognition are *cognitive mechanisms* of some sort. He clearly distinguishes his position from Chomsky's whom he classifies as an advocate of *neocartesianism*. Neocartesianism, Fodor explains, is a view according to which what's innate is "a certain *body of information*" (Fodor, 1983, p. 4). On Fodor's reading of Chomsky,

Q143. When Chomsky says that there is an innately specified 'language organ,' what he means is primarily that there are truths (about the structure of possible first languages) that human beings innately grasp. When he says that the mind of the child is 'intrinsically structured,' what he means is primarily that there are innately specified propositional contents. (Fodor, 1983, p. 7)

Fodor does not find Chomsky's view very plausible. He explains:

Q144. 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. 5–6)
A similar reasoning, I believe, could be applied to Fodor's views on the innate databases of modules, and more generally to his account of concept acquisition. As we said above, the account of both transducers and input systems certainly requires certain 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 the output that they produce. It just doesn't make sense to claim that these faculties will be able to do their job only if what they produce as their output are innate mental representations. There is no need to posit either innate percepts, or innate concepts. Rather, I think that as it was the case with Aquinas, so it is here. According to Aquinas, our sensory organs only register discreet physical data. However, we don't see sensible species, but rather substances, or objects. That's only possible because of how our mind works: because our mind includes the internal senses, that is, innate internal mechanisms, whose function it is to transform external stimulation into the format that our brain will be able to use as representations of things in the world. We don't know how exactly the mind does it, but this mystery does not imply that the representations that come up at the output of the process of cognition have to be innate.

Consider Fodor's description of the faculty of memory.

Q145. If one is going to postulate innately specified faculties, memory is, surely, a plausible candidate. Yet (…) having a memory isn't a matter of having one or another set of beliefs, and if memory is an innate capacity, that couldn't be because there is some set of propositions that organisms are born cognizing. (…) Memory is, so one supposes, some sort of mechanism, analogous to a hand or a liver or a heart. (Fodor, 1983, p. 8)
As Fodor himself puts it (while criticizing Chomsky), what we need is not some kind of "an 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

Q146. 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)

To sum up, neither Fodor's understanding of transducers, nor his account of input systems imply or require propositional content, or innate concepts of any kind. Fodor's criticism of Chomsky indicates that innateness of modules is architectural. The functioning of modules may also imply organism nativism in the sense that the kind of modules that we have determines what kinds of mental representations we will ever acquire. There does not seem to be any reason, however, to believe that the modularity of the mind implies or requires Neural Nativism.

iii. Implications of Fodor's Account of Concepts

There is a passage in Fodor's 2008 which sounds very promising from the point of view of this dissertation. The conclusion that Fodor expresses in this passage seems quite close to the view which, based on our considerations, it seems that he should adopt. He says:
Q147. The central issue isn't which concepts are innate, (...) [because] there are none of those (...). Rather, the problem is to explain how a creature's innate endowment (...) contributes to the acquisition of its conceptual repertoire; that is, how innate endowments contribute to the processes that start with experience and end in concept possession. (Fodor, 2008, p. 145)

Unfortunately, just a few lines later Fodor goes back to his favorite mode of thinking and announces that "it's time to think about what a nativist theory of concept acquisition might be like" (Fodor, 2008, p. 145). So, in the end, in his LOT2 yet again he takes it to be a 'common ground principle' (CG1) that we must have some innate concepts. The CG1 states:

Q148. Minds like ours start out with an innate inventory of concepts, of which there are more than none but not more than finitely many. (Fodor, 2008, p. 131)

As we have seen, nothing in Fodor's account of the architecture of the mind indicates that he is justified in supporting this claim. Similarly, we won't find support for Fodor's representational mad-dog concept nativism in his account of concepts as thought-parts.

Recall that in virtue of his physicalism, Fodor thinks that concepts, which are parts of sentences of LOT, 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 (see above, 4.1.2).

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 to accept Neural Nativism. What Fodor would need is not only
the kind of tissue nativism which negates multible realizability, but also the kind of
nativism which would suggest that 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 Fodor would need in order to claim 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 does not seem plausible from a scientific point of
view (5.3.5). Identity between individual concepts and specific neural patterns in specific
locations in the brain would not be an evolutionary adaptation. What I want to suggest
is that Fodor himself does not need to support such a position. Consider his diagnosis to
the effect that the argument about innateness was really about

Q149. some one thing—after all: it was about whether there is a characteristic human
psychological phenotype (...) that can be attributed to a characteristic human genetic
endowment. The constellation of notions that cluster around 'genetic determination' 'genome'
'genotype' and the rest are, to be sure, themselves adequately contentious. We have nothing
like a general account of how genotypic variables are phenotypically expressed, and it's clear
that epigenetic and environmental processes are, practically invariably, interacting factors.
Understanding how such interactions work may well be the hardest problem that science has
ever faced. But the idea that some phenotypic properties are significantly genotypically
determined is by now deeply scientifically entrenched; to that extent, biology appears to be in
the process of constructing a concept of innateness that saves many of the rationalists'
paradigms. Skin color really is largely genetically determined, hence significantly heritable
(...). Likewise birdsong in a lot of cases; likewise the Babinsky Reflex. (...) We have, in short,
good reason to take for granted that there's a substantive notion of innateness because
biology needs one however the rationalism/empiricism issue turns out. (Fodor, 2001, p. 102)

The passage suggests that Fodor should be inclined to accept not a pessimistic, but rather
a biological notion of innateness. If he'd accept a definition of innateness similar to the
one that I propose, he'd 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 the 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. 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 the 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's innate in all this is the complex functioning of various cognitive mechanisms.

Consider yet another passage from Fodor:

Q150. An organism's mental-state tokens get caused by, for example, events that transpire in the organism's local environment. There are, of course, mechanisms—typically neuronal ones—that mediate these causal transactions. And these mechanisms have presumably got an evolutionary history. They are presumably the products of processes of selection, and it's not implausible that what they were selected for is precisely their role in mediating the tokening of mental states. So there are these cognitive mechanisms, and there are these cognitive states; and the function of the former is to produce instances of the latter upon environmentally appropriate occasions. (…)

Strictly speaking it doesn't, of course, follow, that the cognitive states themselves (…) have any function at all. (…) Since the assumption that there is a teleological story to be told about the mechanisms of belief fixation does not imply that there is a teleological story to be told about
beliefs, it a fortiori does not imply that beliefs (or, mutatis mutandis, other intentional states) can be individuated by reference to their functions. This is important because it's more intuitive that belief-fixing mechanisms (nervous system, for example) have functions than that beliefs do; and the implausibility of the latter idea ought not to prejudice the plausibility of the former. (Fodor, 1990, p. 65)

As it seems, therefore, 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).

iv. Fodor and Aquinas: Final Remarks

We have shown that Fodor and Aquinas share (or plausibly could share) the following claims with respect to cognition:

There are three levels of cognition.

In the first stage of cognition, transducers, or external senses, register stimuli coming from the external world and translate information that they receive into a code that will be understandable to higher-level cognitive faculties.
In the second stage of cognition, input systems, or the internal senses (especially the cogitative power), take as input the output of the transducers/external senses and produce representations of external objects. Concepts are produced in the second stage of cognition. Concepts are thought-parts. An innate trait can be defined as an evolutionary adaptation, genetically inherited within a given species. There are various architectural innate elements in the human cognitive endowment.

(A) Are concepts innate?
For Fodor, the role of input systems is to infer the arrangement of local distal objects. That's why he believes that input systems include perception and language. The outcome of perception and language are percepts representing objects in the external world, and concepts, that is, mental symbols allowing us to think about those objects. (Fodor seems to be talking about language in the sense of speech recognition and understanding; but because of the relationship that he thinks holds between any natural language and the language of thought, it seems plausible to talk about concepts in this context.) On Aquinas's account, the cogitative power which collaborates with the intellect in producing concepts, also belongs to the level of the senses responsible for perception. This suggests that both Fodor and Aquinas could conclude that concepts are as innate as percepts are. But this means that concepts are not innate, only the mechanisms that produce them are. Neither philosopher needs to posit innate concepts or innate representations. As we have seen, the view according to which concepts are innate is actually scientifically implausible.
The remaining task is to consider the results of our considerations for the transduction and the d/D problems.

(B) The Transduction problem

Fodor's input systems, as we said, are supposed to be 'computational mechanisms' whose role is to produce representations of distal objects. The question that needs to be answered, however, is whether computations understood as inference-like processes, which take premises and produce conclusions, can indeed produce representations of external objects. Fodor himself seems to be aware of the problem. He says:

Q151. Input systems function to interpret transduced information and to make it available to central processes; (...) in the normal case, what they provide will be information about the 'layout' (...) of distal stimuli. How might such a system work? Heaven knows there are few harder questions; but I assume that, in the case of perception, the answer must include some such story as the following. The character of transducer outputs is determined, in some lawful way, by the character of impinging energy at the transducer surface; and the character of the energy at the transducer surface is itself lawfully determined by the character of distal layout. Because there are regularities of this latter sort, it is possible to infer properties of the distal layout from corresponding properties of the transducer output. (Fodor, 1983, p. 45)

Instead of an explanation of what exactly such inferences would consist in, or what would be their origin, Fodor offers an example:

Q152. A useful example is Ullman's (1979) algorithm for inferring 'form from motion' in visual perception. Under assumptions (e.g., of rigidity) that distal stimuli usually satisfy, a specific sequence of transformations of the energy distributions at the retina will be reliably interpretable as having been caused by (and hence as specifying) the spatial displacement of a distal object of determinate three-dimensional shape. A device that has access to the transducer outputs can infer this shape by executing Ullman's (or some equivalent) algorithm. (Fodor, 1983, p. 45)
Fodor assumes that "performing such computations is precisely the function of input systems" (Fodor, 1983, p. 45).

What Fodor is saying here is that—in this case, visual—input systems 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, then, we want to consider is the issue of concept acquisition, it turns out that what we need to posit are also certain innate mechanisms that apply the right kinds of algorithms to the data received and as a result produce mental symbols which represent the distal objects, or patterns of neural activations whose firing constitutes thinking a given thought-part. What Fodor does not say is where such algorithms come from. It seems that the only plausible answer to this question is that the functioning of our cognitive modules 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 mechanism, that is, input systems, that produce (usually in a reliable way) representations of distal objects. The representations themselves need not be innate.

This appears similar to Aquinas's solution. Aquinas could not solve the transduction problem if he were to remain the kind of FT-empiricist that many of his interpreters would want him to be. At the same time, what he needs to solve the problem of the passage from the sensible to intelligible species are not innate concepts. Rather, Aquinas's solution to the transduction problem is to be found in his idea of the
cogitative power, an innate cognitive mechanism, collaborating with the intellect. The
cogitative power can be interpreted as an evolutionary adaptation—a mechanism that is
advantageous to the species' survival. It is a fixed innate trait for human beings, chosen
by natural selection because of its advantage for the species, that based on the
information provided by the external senses they are able to cognize individual
substances.

(C) The d/D problem

The discussion regarding the transduction problem already indicates a solution to the
other problem that both Fodor's and Aquinas's accounts of cognition face, the d/D
problem. In his considerations regarding the d/D problem Fodor is worried about
randomness of the (causal) relationship between experience and resulting concepts. It
seems, however, that he should be equally worried about randomness of another causal
relationship: the relationship between experience and resulting *percepts*.

Fodor could have had a significant advantage over Aquinas's theory if the
hypothesis testing method worked as an explanation of concept acquisition.
Unfortunately, this is not the case. More importantly, the hypothesis testing method
itself faces the d/D problem. And, therefore, Fodor's advantage over Aquinas
evaporates.
BIBLIOGRAPHY


------ (2001). *Four ages of understanding. The first postmodern survey of philosophy from ancient times to the turn of the twenty-first century*. Toronto: University of Toronto Press.


----- (1998b). There are no recognitional concepts; not even RED. In E. Villanueva (Ed.), *Concepts: philosophical issues*, 9 (pp. 1–14). Atascadero, CA: Ridgeview.


------ (1997b), *Theories of cognition in the later Middle Ages*. Cambridge: Cambridge University Press.


