# Massive Modularity:
# Why it is wrong, and what it can teach us anyway

by

Drew Blackmore

A thesis submitted to the Faculty of Graduate and Postdoctoral
Affairs in partial fulfillment of the requirements
for the degree of

Master of Arts

in

Philosophy

Carleton University
Ottawa, Ontario

## Abstract

This thesis addresses current issues of cognitive architecture, with a focus on the family of theories known as massive modularity. This project begins with a discussion of the concept of modularity as proposed by Jerry Fodor. All of Fodor's criteria of modularity are explored in order to establish a formal definition of modularity. That definition is then used as a benchmark to determine whether the cognitive mechanisms proposed in the massive modularity theories of Leda Cosmides, John Tooby, Dan Sperber, Steven Pinker, and Peter Carruthers actually qualify as modules. After concluding that the massive modularity theories of the above authors are in fact not modular, the discussion turns to Zenon Pylyshyn's cognitive impenetrability thesis in order to demonstrate that it is extremely unlikely that there could exist any truly modular version of the massive modularity hypothesis. Finally, an alternative account of the mind is proposed in place of massive modularity.

## Acknowledgements

I would like to thank my supervisor, Dr. Andrew Brook, for his infinite patience and invaluable critique over the course of this entire project. I would like to thank Dr. Heidi Maibom, Dr. David Matheson, and Dr. Jordan Dodd for all their wonderfully helpful and educational comments at almost every stage of this piece. I would like to thank my parents for their help in editing, and I would like to thank my partner, Emily Griggs, for her constant support. Finally, I would like to thank Nadia Kharyati and all the employees at the Raw Sugar Cafe for providing the best writing atmosphere I ever could have hoped for.

Table of Contents

## Introduction

Over just the past few decades, it seems that the fields of psychology and biology have been growing closer and closer. Largely, this appears to be because the belief that "the human brain did not fall out of the sky, an inscrutable artifact of unknown origin," is now prevalent (Cosmides & Tooby, 1994, 85). Put more seriously, it is now the case that, excluding those with a vested interest in denying evolutionary theory, most agree that the human mind evolved. While it is only natural that this discovery is more important to some branches of cognitive theory than others, it has almost completely turned the field of cognitive architecture in particular on its head. By adopting the theoretical commitments of the evolutionary approach, psychologists' old architectural theories have come under scrutiny by new analytical methods. Perhaps most importantly, this has resulted in decisive arguments in favour of dropping several old, foundational assumptions.

Chief among these bygone tenets is the distinctly empiricist idea that the human mind is born a blank slate, a *tabula rasa*, or an equipotential, domain general, content-independent computational machine. Expressed perhaps most purely by the associationists (Fodor, 1983, 29), but held onto by other thinkers of a far less radical bent, the main idea behind the domain general approach concerns innate knowledge and the structure of the faculties of the mind. For the mind to be equipotential, domain general, and content-independent, it must possess effectively no innate knowledge, and be essentially structured entirely out of mechanisms that can tackle any and every problem, situation, input, or computation with equal capacity. In this view, the final shape of the mind is defined primarily by external factors, such as culture, which are considered sufficiently powerful to replace innate dispositions (tendencies, proclivities, drives, etc.). Thus, things like sexual orientation, personality type, or even learning procedures are thought of as ultimately determined by external influences. In short, it does not matter "if the clay of the human mind has some initial shape…, so long as it is soft enough to be pounded by the external forces into any new shape required" (Tooby & Cosmides, 1992, 29). Essentially, the domain general view holds that the mind is a single general-purpose, information-processing machine (or computer) that brings to bear minimal innate knowledge about the world, and which develops its capacities through repeated interactions with external inputs (or through general learning).

This domain general approach to cognitive architecture began to come under fire not so long ago, when Chomsky developed a new approach to evaluating existing psychological theories of language and language acquisition using solvability analyses. By testing the ability of hypothesized computational mechanisms to perform tasks or solve problems we already knew that the human mind was capable of performing and solving, Chomsky proved a variety of explanations for human linguistic competence to be false, including essentially every variety of general-purpose computational mechanism we knew of at the time (Tooby & Cosmides, 1992, 95). Thus was introduced the concept of 'mental organs,' and the idea that the mind was innately structured as a rich and diverse host of functionally individuated capacities or faculties, in a way similar to the body. This view left behind the idea that the mind was a primarily undifferentiated, general-purpose computer that underwent uniform development across every cognitive domain from perception to memory to judgement (Fodor, 1983, 3). Decidedly not empiricist, Chomsky's theory was that the mind consisted of many bodies of innate knowledge specific to various informational domains, such as language, numbers, or music, and that the computational mechanisms of the mind employed that innate knowledge in order to successfully perform their complex functions.

It was not too long, however, before Fodor demonstrated that Chomsky's picture only went part of the way to denying the domain general account of cognition. Insofar as the domain general account posits *both* that there is no innate knowledge in the mind, and that the computational mechanisms of the mind are equipotential and general-purpose across informational domains, Chomsky's theory only half-refuted it. He posited the existence of innate bodies of knowledge, but still held that primarily general-purpose computational mechanisms realized behavior, simply employing those innate bodies of knowledge in order to do so (Fodor, 1983, 7). Enter Fodor's concept of modularity, an architectural notion meant to describe a certain variety of cognitive mechanism. Fodor held that the mind was divided into two main portions: central cognition, which included such capacities as conscious judgement and analogical reasoning; and peripheral cognition, which was exhausted by the capacities of sense perception and language (Fodor, 1983, 102-103). In order to substantiate this sharp divide, Fodor appealed to the architecture of the capacities that compose peripheral cognition, claiming that peripheral architecture (and peripheral architecture alone) is universally modular.

Despite the name, Fodor's modules were a much closer analogue to bodily organs than Chomsky's 'mental organs' ever had been. Fodor conceived of modules as independent mechanisms in the mind, virtual machines that worked in isolation from one another. Just as how a heart could continue to pump blood even in absence of a liver to purge that blood of poisons, modules were meant to be fully dissociable components that could fully perform their functions even in the absence of the remainder of cognition. Furthermore, these 'snap-in' components which realized perception and language were thought to be innately specified; genes were thought to code for the various regions of the brain that made up these components, meaning that every human's perceptual and linguistic faculties would develop roughly the same way (Fodor, 1983, 100). Thus, Fodor's theory, at least in a portion of the mind, completely denied the domain general picture of cognition. Modules came equipped with the same innate knowledge Chomsky had proposed, while at the same time, each and every module was a domain specific computational device, working with and processing data only from a single informational domain (such as colours or shapes for vision, and pitch or volume for audition). Modules were not equipotential, general-purpose computers, but instead dedicated experts capable only of working with certain kinds of data (Fodor, 1983, 47). All of this made the concept of modularity the antithesis of the domain general approach.

Remember, however, that Fodor's picture of the mind was a hybrid one. Fodor considered only peripheral cognition to be modular. When it came to central processes, Fodor maintained that they were domain general in nature. His view of the mind, therefore, consisted of a single domain general, central processor flanked by a series of perceptual and linguistic modules. The purpose of

Fodor's central processor was to collect and collate the outputs of the surrounding modules, and then generate beliefs based upon the many different kinds of output those modular systems provided. He spent the latter portion *The Modularity of Mind* arguing for the impossibility of central modularity, claiming that the independent, dissociable nature of modules made them incapable of the more global variety of thought involved in analogical reasoning and judgement. In short, Fodor held that if modules are separated from one another, and thus cannot communicate, then comparisons and comparative inferences across informational domains could not even occur. As Fodor claimed,

> what's known about the flow of water gets borrowed to model the flow of electricity; what's known about the structure of the solar system gets borrowed to model the structure of the atom; what's known about the behavior of the market gets borrowed to model the process of natural selection, which in turn gets borrowed to model the shaping of operant responses (Fodor, 1983, 107).

The idea here is that, in order for modules to be fully independent from one another, they need to be 'informationally encapsulated,' meaning that they do not have access to one another's data and outputs (they cannot communicate). Thus, since each module is specific to a single informational domain, a fully modular mind would be incapable of drawing comparisons, or in general using any information from one domain to influence information from another. As Fodor's examples show, however, we evidently are capable of drawing comparisons and using information from one domain to influence information from another. Thus, he claimed, the mind cannot be fully modular (Fodor, 1983, 107).

Perhaps because Fodor thought that scientifically investigating and understanding central cognitive processes was ultimately impossible due to their global nature (Fodor, 1983, 38), other thinkers were relatively quick to resist or reject his domain general picture of central cognition. Since modularity seemed the only way to do so, however, such thinkers immediately latched onto what Fodor had previously called "modularity theory gone mad" (Fodor, 1987, 27), positing that indeed the whole mind was structured out of modules. Somewhere along the line, the theory was christened massive modularity (likely because a massive number of modules would seem to be necessary in order to realize every human cognitive capacity), and this new, entirely domain specific paradigm of cognitive architecture became a major player on the scene. At first, massive modularity was largely the province of evolutionary psychology, a research platform which held that, since humanity had spent almost all of its evolutionary history as hunter-gatherers, all uniquely human cognitive capacities were likely to be evolved adaptations to such a lifestyle (Cosmides and Tooby, 1994, 87; Sperber, 1994, 42; Pinker, 1997, 21). Eventually, massive modularity was also adopted by Peter Carruthers, who gave it what many consider to be its most sophisticated defense thus far articulated.

As might be obvious given the title of this work, I believe that massive modularity is a doomed enterprise. Furthermore, I believe that the approach taken by massive modularists to the concept of modularity is largely what is to blame. In attempting to reasonably cast the whole of the human mind as modular, thinkers have watered down the concept of modularity, so to speak, creating a weaker version of it in order to protect themselves against counterarguments similar to Fodor's original comparative reasoning objection. Most importantly, these thinkers have downplayed the importance of informational encapsulation to modularity; as a result, many 'modules' proposed by these thinkers are not informationally encapsulated even in the slightest. In short, I believe that by doing this, proponents of massive modularity have created a whole variety of theories, but that none of these theories are truly modular. Recognizing this is important, not only because in doing so we clear up a significant amount of conceptual confusion, but also because it opens up the floor to other domain specific views of cognition. As I hope to demonstrate over the course of this project, domain specificity and modularity do not necessarily go hand-in-hand. Therefore, understanding exactly why the theories of various massive modularists are not modular will ultimately open the doors for new, more subtle conceptions of cognitive architecture.

Thus, the purpose of my project could be understood to be twofold. First, I aim to follow the above line of argument until its very end, demonstrating both that every current version of massive modularity, from the evolutionary psychologists' to Carruthers', is endemically flawed, and that there cannot be a version of massive modularity that does not suffer from this same endemic flaw. Since all massive modularists reject informational encapsulation (and indeed must, as I aim to show), they simply cannot establish the modular status of the mechanisms they propose. Put another way, I aim to show that massive modularity, by its very nature, cannot actually be modular at all. Second, I aim to disentangle modularity at large from the theories of its proponents. I believe that, within the theories of both Carruthers and the evolutionary psychologists, are a great many valuable insights, many of which seem to have been obscured or ignored *simply because of the highly controversial modular baggage these theories carry with them*. By separating these insights from their foundations in massive modularity, I hope to propose my own domain specific picture of the mind that has many of the advantages of a massively modular approach, without suffering from its weaknesses.

This project will begin with the foundations of modularity as a concept. While Fodor did not originally define exactly what it means to be a module, he nevertheless offered a variety of criteria that a given mechanism must (mostly) meet in order to be considered a module. My first chapter explores those criteria in detail, focusing on those which are most important to the concept of modularity, as well as the various logical relationships among them. After laying out that groundwork, I will explore the massive modularity theory espoused by evolutionary psychologists Leda Cosmides, John Tooby, Dan Sperber, and Steven Pinker. By applying the work done on the criteria in the first chapter to their version of massive modularity, I will demonstrate that it is, in fact, not truly modular at all, insofar as all the 'modular' mechanisms they propose fail to meet enough important criteria to be properly captured by the concept. Following this, my third chapter will tackle the massive modularity theory of Peter Carruthers, as it appears in his book *The Architecture of the Mind*. Once again, by applying the work done in chapter 1 on the criteria (though in a different manner), I will ultimately demonstrate that even Carruthers' very sophisticated version of massive modularity fails to be truly modular.

In my fourth chapter, I hope to hammer the final nails into the coffin of massive modularity by demonstrating that the single criterion most important to modularity, informational encapsulation, simply is not a prolific property of the mind. I will focus on the cognitive impenetrability thesis of Zenon Pylyshyn, an idea commonly considered to be the psychological mechanism underlying informational encapsulation, bringing to bear a platform of counterexamples and arguments in order to demonstrate that it is likely false for all but a small portion of the mind, and that therefore informational encapsulation is rare. Finally, in my fifth chapter, I will provide an alternative picture of the mind – one which is primarily domain specific, but not massively modular. This account will appropriate many arguments and ideas from the various authors I will discuss in earlier chapters, but will combine those ideas with an approach to functionally individuated domain specific mechanisms proposed by Vincent Bergeron. By appealing to Bergeron's dissection of the notions of function and specificity, I hope to demonstrate the plausibility of this alternative account. Ultimately, I hope to provide reason to investigate the possible truth of a primarily domain specific, but non-modular picture of the mind.

## Chapter 1 – What it means to be modular

When Fodor first introduced the concept of a module in his 1983 work, *Modularity of Mind: An Essay on Faculty Psychology*, he did so in an unfortunately vague way. Exactly whether or not this was a mistake on Fodor's part, however, does not really matter. What is important is clearing up the confusion that followed, ideally by providing a proper definition of modularity in one way or another. Doing just that is my goal in this first chapter. Where Fodor originally listed nine criteria of modularity, essentially claiming that if a mechanism meets most of them, it is likely to meet all of them (Fodor, 1983, 38), I hope to provide at least as exact a definition as one might hope for, articulated in terms of necessary, sufficient, and distinctive conditions. I will begin by explaining each of Fodor's nine criteria. I will then discuss the various logical connections between Fodor's criteria, as I believe that understanding this complex logical web is integral to understanding the concept of modularity. Finally, I will finish this chapter with an argument to the following effect: one of Fodor's nine criteria, namely informational encapsulation, is so important to the concept of modularity as to be essentially equivalent to it, and therefore any mechanism that is not informationally encapsulated is not modular.

Fodor's criteria themselves are as follows: (1) mandatory operation, (2) shallow outputs, (3) fast operation, (4) dissociability, (5) fixed neural architecture, (6) characteristic ontogenesis, (7) informational encapsulation, (8) inaccessibility, and (9) domain specificity. When Fodor originally identified these nine criteria, he claimed that all cognitive mechanisms we can classify as modular should meet the criteria, as a group, at least to some interesting extent. He also claimed that both the concept of modularity and many of its criteria admit of degrees (Fodor, 1983, 37-38). This means that a cognitive mechanism does not need to meet all the criteria in order to be considered a module, as long as it meets at least some them in a strong way.

Section 1 – An introduction to Fodor's nine criteria

At its core, mandatory operation (1) concerns whether a given cognitive mechanism is in any way controlled by the remainder of cognition. In this way, it breaks down into two salient levels: uncontrolled by consciousness and uncontrolled by the rest of cognition at large. If a mechanism is uncontrolled by consciousness, which is to say that it is capable of both initiating and ceasing operation without any conscious effort whatsoever, then it is automatic (Bargh & Chartrand, 1999, 463). If a mechanism is uncontrolled by all the rest of cognition, then whether it initiates or ceases its operations is entirely dependent on its own processes. To illustrate, one cannot help but hear words (in a language one knows) as words, as opposed to noise, even if it would currently be in one's interest not to. Even if hearing those words as words would offend us, confuse us, or perhaps even place us in grave danger, we will hear them as words nonetheless. This suggests that such linguistic parsing is entirely self-controlled, as otherwise some other mechanism would likely deactivate it in such situations (Fodor, 1983, 52-53).

Mandatory operation (1) is also generally seen to be related with shallow outputs (2) and fast operation (3). Delivering shallow outputs (2) boils down to delivering outputs that can be inferred from their corresponding inputs without the aid of any information not contained in what is transduced. This means that shallow outputs cannot be conceptual *qua* unobservable or theory-laden; they cannot be the products of non-deductive inferences from their inputs, and they cannot consist of entities that are not present in the world around us (such as social constructs). Fast operation (3) is essentially near-instantaneous computation; maximally accurate visual processing takes less than half a second, and linguistic parsing seems to happen as fast as is possible, given the constraints of human language (Fodor, 1983, 61-62).

Dissociability (4), a notion Fodor pointed toward in claiming that modules demonstrate characteristic breakdown patterns (Fodor, 1983, 99), concerns the interactions between mechanisms with regard to function and damage. If mechanism M is dissociable, then damage to M would impair M's function without impairing any other cognitive functions. Similarly, damage to those other functions would not impair mechanism M (Fodor, 1983, 99-100). This can be understood in two ways, one structural and the other functional, both of which will receive more detailed attention later. A similar dual understanding goes for fixed neural architecture (5). One way of conceiving of this criterion is in terms of enforced physiological separation between the circuitry of a given mechanism and the remainder of cognition, thus fixing the neural architecture of the mechanism. Alternatively, fixed neural architecture (5) could simply amount to innateness (viz., neural architecture that is fixed insofar as it is roughly identical across individuals of the same species). Innateness also suggests a characteristic (or innately specified) ontogenetic development (6) of the mechanism; in short, the idea is that recreating the exact developmental pathway necessary to create such similar architecture across effectively all individuals of a single species would be so complicated that it could only be the result of genetic specification (Cosmides and Tooby, 1994, 89-90).

Informational encapsulation (7) concerns what information a cognitive mechanism has access to. If mechanism M is encapsulated, then its internal operations cannot make use of any specific set of information outside of either its input or its own dedicated databank (if it even has one). This is not a measure of amount, but of kinds; a specific set of information merely means the information used by a specific set of other mechanisms (including sets of one). Thus, M is encapsulated with regard to mechanism X just in case M's internal operations cannot make use of any information stored/outputted by X. Note that the distinction here between not having access to information and having access to but not being able to use that information is immaterial; in effect, these two states are equivalent. This will become particularly important later on, as it means that for a given mechanism M, if M cannot make use of some set of information S, then M is *effectively* encapsulated with regard to S, even if M is not *genuinely* restricted in its access to S.

Inaccessibility (8) concerns how other mechanisms interact with M. In order to be inaccessible, the internal operations of M must be opaque to other mechanisms; only the outputs of

M, then, are available to other mechanisms. Finally, there is domain specificity (9): in order for M to be a domain specific mechanism, it must be capable of processing information pertaining only to a particular subject matter. Put another way, mechanism M is domain specific just in case it can compute only over information of a limited number of kinds (Samuels, 2006, 50-51). Note that this definition of domain specificity runs contrary to that of Max Coltheart and others of the same persuasion, who hold that mechanism M is domain specific just in case its processes are triggered (or activated) only by inputs that are of a single informational kind (Coltheart, 1999, 119). Given that this interpretation is a departure from Fodor's original conception of domain specificity, I will not be addressing it in my attempt to construct a definition of modularity. This choice is especially important given the fact that these last three criteria of informational encapsulation, inaccessibility, and domain specificity are commonly understood to be the criteria which are most central to the concept of modularity. Fodor himself focused on encapsulation specifically as the hallmark of modularity (Fodor, 1983, 71), which I believe to be correct. As we now move on to a discussion of how the nine criteria relate to one another, I hope my reasons for doing so will become clear.

Section 2 – Logical relationships: Criteria (1) through (4)

We will begin the discussion of the logical connections between Fodor's nine criteria with mandatory operation (1), shallow outputs (2), and fast operation (3). It seems to me that at least the former two are implied by encapsulation (7) in one way or another. In seeing how this is the case for mandatory operation (1), consider the following. If a mechanism cannot make use of *any* information outside of the mechanism itself, then it necessarily must both initiate and cease computation independent of all external cognitive influences. Put another way, if the processes responsible for starting and stopping the mechanism's computations did not operate exclusively on their own accord, they would have to be making use of at least some information outside the mechanism running them. Thus, if a mechanism is completely encapsulated, it must have mandatory operations; it may very well be mandatory despite being more weakly encapsulated, but we can only come to know for sure through empirical research. As such, the strongest form of informational encapsulation (7) necessarily implies mandatory operation (1). If we know that a given mechanism is encapsulated with regard to consciousness, however, then we know that it is automatic, as opposed to fully mandatory. Since such a mechanism would be capable of making use of information from other cognitive mechanisms not involved in the realization of consciousness, its operations might initiate or cease at the behest of those other mechanisms. Thus, while we could not intentionally engage or disengage such a mechanism, other subconscious processes in our minds might be capable of doing so.

If a mechanism's internal operations were incapable of making use of conceptual information, then said mechanism's outputs would necessarily be shallow (2). This too seems to be encapsulation, but encapsulation *with regard to which mechanisms*? There are likely many mechanisms that perform operations on data that concerns conceptual entities. As such, being encapsulated from a single mechanism is not enough to be incapable of making use of conceptual information. The answer is simply encapsulation with regard to *all* mechanisms that store/output conceptual (unobservable, or theory-laden) data. Thus, shallow outputs (2) do not follow directly from encapsulation (7) *simpliciter*, but instead for a specific variety thereof.

As for speed (3), while it is not implied by encapsulation, it is related to it. Consider: if mechanism M integrates less information than its neighbor, it takes on less of a computational burden; assuming both M and its neighbor have the same processing power, M will nevertheless perform its computations faster. This is because encapsulation is a means of achieving computational frugality. The more encapsulated a mechanism is, the more frugal it is, and the more frugal it is, the faster it is likely to be (Carruthers, 2006, 53). Thus, encapsulation leads to fast operations (3) insofar as speed and computational frugality tend to go hand-in-hand.

As stated earlier, dissociability (4) can be understood in two ways. When understood structurally, it has little to do with modularity, as it is simply the realization of different functional mechanisms in non-overlapping neuronal assemblies. As such, it neither implies nor is implied by other criteria of modularity. The various subsystems of a computer might exist on separate chips or boards, but that says nothing as to whether or not they are functionally independent, fast, mandatory, inaccessible, etc. When dissociability (4) is understood functionally, however, it is indeed implied by informational encapsulation (7). Consider this: if mechanism M does not make use of any information stored in/outputted by its neighboring mechanism X, then it will suffer *absolutely no* failure in functionality when X does. If M *both can and does* make use of some information stored in/outputted by X, X's failure would leave M without all the information it needs to complete its own operations, leading to a failure in M's functionality as well. At the same time, if X makes use of any information either stored in or outputted by M, then functional failures of M will cause functional failures in X. As such, M's informational encapsulation with regard to X implies M's dissociability from X, and *vice versa*. This also means that degree of encapsulation is directly related to degree of dissociability; if M is informationally encapsulated from all of X, Y, and Z, then M is dissociable from all of them as well. The more mechanisms M is encapsulated from, the more dissociable M is.

Despite how related these two criteria are, dissociability (4) does not imply informational encapsulation (7). Put simply, this is the case because informational encapsulation is not the only way to achieve dissociability. If mechanism M is only dissociable from its neighbor X, this could very well be the case without M being informationally encapsulated from X. This is because, as I mentioned in the last paragraph, in order for X's failure to leave M without all the information it needs to complete its own operations, M both *must be capable of* making use of some information stored in/outputted by X, and *must actually do so*. If M has access to the information in X and can use it, but simply never does, then M will not suffer failures in functionality when X does. Think of this as a form of *effective* functional dissociability, as opposed to a hard-wired or hard-coded, *genuine* functional dissociability. Thus, the mere fact that a theory of mind mechanism has been seen to be dissociable from a face recognition mechanism, for example, does not prove that the former is informationally encapsulated from the latter. It could be that the theory of mind mechanism has full access to the facial recognition mechanism, and can make use of the information the latter stores/outputs, but simply does not. If we knew that the theory of mind mechanism did not even have *the potential* to suffer a loss of functionality as a result of a failure of the facial recognition mechanism, however, then informational encapsulation would be assured. This distinction would be important when discussing the implications of empirical results.

Section 3 – Logical relationships: Criteria (5) through (9)

As for fixed neural architecture (5), we can largely discount it in this discussion. We have learned a fair amount about the brain since Fodor wrote about modularity for the first time in 1983, and one thing we can be sure of now is that the dedicated circuits that qualify mechanisms as having a fixed neural architecture are essentially non-existent across the brain (Koch, 2004, 23 & 59). As such, the only way fixed neural architecture (5) can be of any interest is if it is construed in a different way. Understood as fixed *qua* species-universal (identical across individuals of the same species), this criterion would amount to innateness, and thus would require an innately specified ontogenetic sequence (6) in order to so reliably develop. Some proponents of modularity construe this criterion exclusively in this manner (see Cosmides and Tooby, 1994). Arguments for innateness do not amount to arguments for modularity, however. Just as with structural dissociabililty, knowing that a mechanism is innate tells us nothing as to its functional architecture. An innate mechanism could be fast or slow, inaccessible or transparent, or whatever else as necessary. Innateness and modularity are related topics, but not logically related concepts.

Inaccessibility (8) also can be discounted in the current discussion. While it is, without a doubt, an interesting feature to predicate of any cognitive mechanism, it also has little to do with any of the other criteria of modularity. The inaccessibility of a given mechanism is irrelevant to whether or not that mechanism is encapsulated, fast, mandatory, dissociable, etc. Put another way, whether or not the mind is composed entirely of inaccessible mechanisms should be an important question to cognitive scientists and philosophers of mind alike, but providing an answer to this question is not the same as providing an answer to the question of whether or not the mind is composed entirely of modular mechanisms.

Domain specificity (9) is another criterion that can be understood in two ways, and how one does so is very salient for any discussion of Massive Modularity. Insofar as domain specificity concerns the kinds of information over which a given mechanism can compute, this computation could be construed with regard to either input or output. If a given mechanism can only take as input a limited number of information kinds, we could call it domain specific with regard to input, or simply input specific (9a). As an example, a hypothetical mechanism involved in vision might only process colour data, and be unable to process any other kind of information in fulfilling its function. Conversely, if a given mechanism can take an unlimited number of information kinds as input, but is limited in the number of information kinds it can output, we could call it domain specific with regard to output. Such a mechanism might be called either output specific or function specific (9b), and I include this second appellation because some authors seem to have already expressed it without making explicit this distinction. Cosmides and Tooby posit a number of examples of function specific mechanisms, such as a social contract mechanism, a cheater-detection mechanism, and an incest-avoidance mechanism. These are not described as domain specific to some set of social inputs, but are yet called domain specific insofar as they perform only a single kind of function (Cosmides and Tooby, 1994, 90).

<u>Section 4 – Why input specificity implies informational encapsulation</u>

At this point, a rather lengthy discussion concerning the logical relationship between domain specificity (9) and informational encapsulation (7) looms before us. Furthermore, since input specificity (9a) and output specificity (9b) are very different beasts, this discussion will need to be split into two parts. Since it is more complex, the discussion on input specificity (9a) will come first. What follows is a simple argument for the implication from input specificity (9a) to informational encapsulation (7). Note that the first premise is merely the definition of input specificity arranged as a conditional statement, and the third premise the same again, but for informational encapsulation. Also, while articulating each premise of this argument as an existential statement would be more technically correct than what I have actually written below, I believe both that the current form is clearer and that it loses no formal efficacy in being so stated.

I. If mechanism M is input specific to (can only process) information of kind φ, then it cannot compute over information of all kinds not-φ.

II. If mechanism M cannot compute over information of all kinds not-φ, then it cannot make use of not-φ information in its processes.

III. If there exists some information that mechanism M cannot make use of in its processes, then it is encapsulated with regard to that information.

C. If mechanism M is input specific to (can only process) information kind φ, then it is encapsulated with regard to information of all kinds not-φ.

This may initially appear perplexing; if informational encapsulation is an access-restriction and input specificity is a processing-restriction (specifically of kind), then how can the latter bring about the former? The answer is two-fold. First, one needs to remember that not having access to information and having access to but being unable to make use of that information are effectively identical. Second, one must carefully consider the distinction between data we call input and data used in processing. Together, these discussions will act as substantiation for premise two, the obvious key premise of my argument above. Immediately below is a demonstration of how the inability to make use of information is *effectively* identical to *genuine* encapsulation.

Consider two cases. In the first case, we have mechanism M, another mechanism A, and the rest of cognition, where all the data stored/outputted by both A and the rest of cognition are available to M (M is not encapsulated at all). In the second case, we once again have M and A, and the same access relation between them, but this time M is informationally encapsulated with regard to the rest of cognition. In both cases, all of A's information that M might access is of kind φ, while all information in the rest of cognition that M might access is of kinds other than φ (read: not-φ). In the first case, M is not input specific (viz., can compute over information of both kind φ and all other kinds); as such, it can process any information from every other mechanism in the entire mind, including A. If we change that, however, and make M input specific to information of kind φ in the first case, then we see a sudden and drastic change: because only A, but not the rest of cognition, stores/outputs information of kind φ, M now cannot make use of any information from mechanisms that are not A. Courtesy of its input specificity, M can no longer process (and thus use) the information stored in the remainder of cognition, and that means that M is unable to compute over it. Thus, if M is input specific in the first case, it can make use of the information in A and itself, but nowhere else. Compare this to the second case, where M's genuine encapsulation means that *it can make use of the information in A and itself, but nowhere else*.

At this point, it should be obvious that making M input specific makes the first and the second cases effectively identical. Despite the fact that an input specific M may not be *genuinely* encapsulated with regard to other mechanisms, M's inability to compute over the information of those mechanisms makes it *effectively* encapsulated with regard to those mechanisms. A theory of mind mechanism specific to the domain of desires, for example, may not be restricted in its access to its neighboring naïve physics mechanism, but the fact that the physics mechanism does not store/output information as desires ultimately makes the theory of mind mechanism incapable of using that information in the exact same way that genuinely restricting its access in the first place would. If a mechanism can speak only English, then it simply cannot be affected by the information in another mechanism that can speak only Greek. In the same way that you or I might be unable to make use of any linguistic information spoken in Japanese (assuming you do not speak the language, of course), input specific mechanisms cannot *understand* information outside of the domain they are specified to. Despite the commonplace language, this is not a homunculus view; all I am claiming here is that being incapable of processing (or understanding) certain kinds of information is effectively the same as not having access to that information. From a standpoint concerned exclusively with the gathering of information, if someone were to give you directions in Japanese, it would be just as if they had given you the directions in English, but you had suddenly gone deaf.

Thus, input specificity (9a) implies informational encapsulation (7), namely a form of effective informational encapsulation with regard to all mechanisms that do not store/output information of the specified domain. Obviously, not having access to a set of information S and not being able to make use of S are two different states of affairs, but what is important here is that they have the same outcome. Interestingly, informational encapsulation (7) does not necessarily imply input specificity (9a). There seems to be nothing stopping a mechanism from being able to process any kind of information, but at the same time being restricted in accessing information from other specific mechanisms; the former does not follow from the latter at all. Thus, this entailment relation from input specificity (9a) to informational encapsulation (7) holds despite the fact that the opposite entailment does not.

Of course, when we use the phrase 'compute over' in the above illustration, it is ambiguous. Does it exclusively mean the data transformed by a mechanism, or does it also include any data involved in that transformation? This is where the distinction between input and processing enters the picture. Consider the following example: a mechanism in the visual cortex perceives the size of various objects, but in doing so it employs its algorithms upon both the visual data it receives from its transducers and some beliefs concerning expected size it acquires from elsewhere in the mind. Both the visual and belief data are data entering the system. Both are involved in the mechanisms' computations in one way or another. In terms of domain specificity, why should one be constitutive of the mechanism's domain while the other is not? The fact that only what we call the input should be so constitutive strikes me as arbitrary. It is indeed the case that the input alone undergoes transformation as a part of the computation, thus marking it as different in a way, but it is also the case that the mechanism must still be able to process (or understand) the additional information that influences those transformations in order for that information to actually influence them. Thus, it seems to me that, while there is a difference between the input data and data used in processing, it is not salient to domain constitution. In short, whether a particular subset of data is transformed or not seems to have no bearing on the fact that the mechanism making use of it still needs to understand it.

This all seems to demonstrate that input specificity (9a) implies informational encapsulation (7). What of output or function specificity (9b), then? Does it imply informational encapsulation as well? The short answer is 'no'. A mechanism is output specific when it can output information only of a single kind. While it is therefore an evidently interesting property to predicate of any cognitive mechanism, output specificity is not interestingly modular (or not a necessary, sufficient, or distinct condition of modularity). We will discuss the argument behind this claim fully in the next chapter, as whether or not output specificity is distinctly modular is more of a concern for evolutionary psychologists specifically than a concern of modularity theorists in general.

Section 5 – A definition of modularity

Now that each criterion has been given its due, what remains is to propose a formal definition of modularity, an exercise that goes hand in hand with arguing for the importance of informational encapsulation to modularity as a concept. Of course, by this point I believe that it is relatively obvious why both Fodor and I consider informational encapsulation (7) to be so central to modularity: it is either logically implied by or itself logically implies so many other criteria of modularity. In their massive modularity theories, evolutionary psychologists place great emphasis on domain specificity (Sperber, 1994, 40; Cosmides & Tooby, 1994, 90; Pinker, 1997, 21). In his book, *The Architecture of the Mind*, Peter Carruthers focuses on functional dissociability (4) as a key property of modules (Carruthers, 2006, 17). Both of these criteria are logically related to informational encapsulation (7), as are many others. Insofar as only informational encapsulation (7) is related to so many important criteria of modularity, it therefore seems obvious that informational encapsulation is the most distinctively modular criterion of the lot.

Furthermore, informational encapsulation is essential to the idea of autonomy that the word 'module' implies (Fodor, 1983, 72-73). When Carruthers discusses modularity, he begins by pointing towards a simple hi-fi system, noting that it is modular insofar as various pieces, such as the speakers or tape-deck, can be removed or replaced without effecting the operations of their neighbors (Carruthers, 2006, 2). When Stephen Pinker addresses massive modularity in his book *How the Mind Works*, he notes that "the word 'module' brings to mind detachable, snap-in components" (Pinker, 1997, 30-31). Central to both these conceptions of modularity is the idea of autonomy, or independence. Evidently, independence has a lot to do with what it means to be modular.

As Fodor points out, informational encapsulation essentially boils down to a form of independence. Even if two mechanisms might compete for access to computational resources, such as memory, they can nevertheless be independent of one another in an informational sense as long as they are informationally encapsulated from one another. Thus, "the question of 'how much autonomy?' is the same question as 'how much constraint on information flow?'" (Fodor, 1983, 73). This is exactly why informational encapsulation (7) implies functional dissociability (4); mechanisms that are encapsulated are also assured in their functional independence. Furthermore, that informational encapsulation *necessarily implies* functional dissociability is why it, and not simply functional dissociability itself, is the central criterion of modularity. As we have already noted, a given mechanism M can be *effectively* dissociable from X by simply not using the information X stores/outputs, even if M both has access to and is capable of using that information. Thus, a given mechanism, if it is dissociable without being informationally encapsulated, can still be massively interconnected with the remainder of the mind. Thus, functional dissociability (4) cannot be central to the concept of modularity, simply because massive interconnectivity is not what the term 'module' implies.

I believe that this is not a question-begging approach. As both Carruthers and Pinker have noted, a module is not simply a functionally independent mechanism – functional independence is distinctive of modularity insofar as only modules are fully functionally independent. Furthermore, functional independence can be achieved only through avoiding massive interconnectedness, and to avoid massive interconnectedness is to be (at least weakly) informationally encapsulated. Put another way, the only way to achieve the distinctive functional independence of modules is through informational encapsulation (either effective or genuine). Therefore, modularity essentially is informational encapsulation (7), not merely functional dissociability (4), even though functional dissociability is, by corollary, an important part of what it means to be modular.

Thus, a proper definition of modularity has presented itself. Informational encapsulation (7) is a necessary, sufficient, and distinctive condition of modularity. It is a necessary condition because, in order to be a module, a mechanism must be informationally encapsulated. It is a sufficient condition because any mechanism that is informationally encapsulated is automatically modular. Informational encapsulation is a distinctive condition of modularity because it is the only jointly necessary and sufficient condition. This means that it is an exclusively modular property; only modular mechanisms are informationally encapsulated, and therefore informational encapsulation is the distinctive feature of modules – to be informationally encapsulated is what it means to be modular. As for the other criteria: at least some degree of functional dissociability (4) is a necessary condition of modularity, insofar as it is implied by informational encapsulation both directly and in degrees. Input specificity (9a) is a sufficient condition of modularity, as it implies informational encapsulation. All the other criteria are related to modularity, but are not important to defining modularity in the same way as those above.

Of course, at this point, you might be concerned that this is really just all terminological. Why can modularity not have a variety of definitions? Why put so much focus on modules as conceived of by Fodor, excluding other conceptions of modularity from the concept? My simple answer: calling functionally non-autonomous mechanisms modules is the type of misnomer that can suggest unintended conceptual baggage, and therefore confuse intellectual endeavor surrounding the concept. Furthermore, insofar as functional independence is distinctive to modularity, calling functional non-autonomous mechanisms modules amounts to loosing that distinctive trait. Despite discussing the hi-fi system as a paradigmatic example of modularity, the modules that Carruthers ultimately proposes are, as we will see in chapter 3, actually nothing like a hi-fi system; they lack the independence that the parts of the hi-fi system have and that the word 'module' suggests. When Pinker discusses massive modularity, not only does he say that "the word 'module' brings to mind detachable, snap-in components," but also that this is misleading, and that "the metaphor of the mental module is a bit clumsy; a better one [to describe his position] is Noam Chomsky's 'mental organ'," (Pinker, 1997, 30-31). Essentially, Pinker is pointing out that the term 'module' suggests a whole galaxy of cognitive properties that he does not want to commit to. This is the very variety of conceptual confusion this project seeks to clarify. There are a great many interesting ideas within the various modularity theories proposed by the authors that I will discuss over the course of this piece, but in order to get at those ideas, they first need to be disentangled from the conceptual baggage they carry with them through their association with modularity. This is why, in moving forward, the definition of ‘module’ I have constructed here will be what I use to determine what is and is not modular. As philosophers of mind and cognitive scientists interested in modularity, by properly defining our terms, we can finally make some headway in determining which architectural theories are and are not viable, and therefore a better idea of exactly what sorts of mechanisms likely make up the mind.

## Chapter 2 – A critique of evolutionary psychology

Now that we have a definition of modularity, it is time to begin our treatment of massive modularity. Given that evolutionary psychologists were the first to embrace massive modularity as a part of their conceptual platform, it seems only appropriate to start our treatment with them. My criticism of the massive modularity of evolutionary psychologists will focus on domain specificity understood in terms of output, of what I have previously called functional specificity (9b). As I promised in chapter 1, section 4, in this chapter I will provide a more complete treatment of functional specificity (9b), focusing on why it is neither a necessary, sufficient, nor distinct condition of modularity. I will then bring these insights to bear upon the massive modularity theories of several prominent evolutionary psychologists, namely Dan Sperber, Leda Cosmides, John Tooby, and Steven Pinker. Ultimately, I hope to show that the ‘modules’ proposed by these thinkers are not in fact modules at all, insofar as none of them meet any sufficient or distinctive conditions of modularity. In so doing, I hope to disentangle massive modularity from what I believe to be the various valuable insights evolutionary psychology can provide – an exercise which I believe will bear fruit later in this project. For now, however, I will focus on demonstrating that the massive modularity of evolutionary psychologists is not truly a modularity hypothesis at all.

Before I launch into my criticisms, however, it seems best to identify the fundamental tenets to which evolutionary psychologists adhere. Despite the fact that Cosmides, Tooby, Sperber, and Pinker each offer different arguments and evidence in favour of massive modularity, all of them share the same adaptationist commitments and focus on investigating the mind *qua* product of natural selection. As Richard Samuels puts it, all evolutionary psychologists view the mind “as composed largely or perhaps even entirely of innate, special-purpose computational mechanisms or 'modules' that have been shaped by natural selection to handle the sorts of recurrent information-processing problems that confronted our hunter-gatherer forebears,” (Samuels, 2000, 13). Despite their differences, it seems to me that all the authors we will discuss in this chapter would agree with

Samuels' summation. All of them posit that the mind consists primarily of modular mechanisms, that those modules are all innate, and that those modules all evolved in order to solve the adaptive problems inherent to a hunter-gatherer lifestyle (Cosmides and Tooby, 1994, 87; Sperber, 1994, 42; Pinker, 1997, 21). In the discussion that follows, I will focus on these commitments, as I believe that they have led evolutionary psychologists to introduce a common brand of massive modularity that relies too heavily on function to distinguish between modules – a strategy which I believe ultimately fails.

Section 1 – Why function specific mechanisms are not modules

When we discussed domain specificity in the previous chapter, I drew a distinction between domain specificity with regard to input and domain specificity with regard to output. I called this latter variety output specificity, or function specificity (9b). To reiterate the idea, domain specificity concerns the kinds of information over which a given mechanism can compute, but this computation can be construed in two ways. Thus, if a given mechanism can take an unlimited number of information kinds as input, but is limited in the number of information kinds it can output, that mechanism is output or function specific. Take as an example a mechanism that could output information only about the colour of perceived objects, even if in so doing it computed over both visual information and belief information about colour expectations. Calling such a mechanism either output or function specific would be fitting, given that any mechanism which can only output information of a single domain would *ipso facto* have a function specific to that domain. This example mechanism's output would therefore be specific to the domain of colour (even though its input would not be), and as such its function would evidently have something to do with colours, such as determining or describing them. Thus, domain specificity with regard to output is function specificity.

In the previous chapter, the fact that input specificity (9a) implied informational encapsulation was one of the key insights. In this chapter, the fact that output or function specificity (9b) *does not* imply informational encapsulation, or any other criterion of modularity for that matter, is absolutely central. Combined with the claim that output specificity is not a necessary, sufficient, or distinct condition of modularity, this means that output specificity has essentially nothing to do with what it means to be modular, which is the basis of my criticism of evolutionary psychology's massive modularity. The criticism is summed up by the following argument:

I. Function specificity is neither a necessary, sufficient, nor distinct condition of modularity alone and does not imply any criterion that is.

II. If function specificity is neither a necessary, sufficient, nor distinct condition of modularity alone and does not imply any criterion that is, then unless the mechanisms described by evolutionary psychologists can meet at least one other criterion that is a necessary, sufficient, or distinct condition of modularity, those mechanisms are not in fact modules.

III. The mechanisms described by Evolutionary Psychologists cannot meet even one other criterion that is a necessary, sufficient, or distinct condition of modularity alone.

C. The mechanisms described by evolutionary psychologists are not in fact modules.

Since failing to propose actual modules means failing to propose an actual modularity thesis, this argument is one that friends of massive modularity ought to care about. If output specificity is not a necessary, sufficient, or distinct condition of modularity, then output specific mechanisms that meet no other criterion of modularity are simply the functionally-individuated workings non-modularists already posit. Therefore, this argument does nothing to invalidate evolutionary psychology at large, but simply demonstrates that the mechanisms proposed by evolutionary psychologists are not modules. In order to substantiate this argument, I will be focusing on premises (I) and (III). As such, you could consider this argument to be an attempt to demonstrate that the merely output specific mechanisms that are the hallmark of evolutionary psychology are insufficiently independent to be modules, as we defined the concept previously, in any meaningful manner.

To begin our treatment of premise (I), then, consider the following: one can conceive of a mechanism that is both function specific and not a module at all. As an example, consider a theoretical mechanism that assesses the social status of other individuals; when engaged, its outputs exclusively concern power relationships between the individual assessed and other individuals either in the surrounding vicinity or in mind at that moment. Such a mechanism seems specific to a single function indeed. However, that this social status mechanism performs this one function only does not immediately and necessarily suggest that this mechanism is a module. This is to say both that the mechanism does not seem to be independent or autonomous in the way the term ‘module’ suggests, and that there seems to be no reason why the mechanism must necessarily meet any of Fodor's criteria for modularity other than its domain specificity with regard to outputs. That our social status mechanism need not be functionally dissociable (4) is obvious. The mere fact that it outputs data only of a single kind says nothing as to whether its computations are autonomous, reliant on one other mechanism, reliant upon several, or even reliant upon bits of data from almost every other mechanism in the mind. The same could be said of any given function; this is because simply knowing that a mechanism performs only a single function (viz., is function specific) gives us no insight into how that mechanism actually goes about performing it.

Function specificity (9b) seems to be separate from more than just independence. That a mechanism is specific to the function of social status (or any function) is a fact that stands apart from whether or not that mechanism meets any of the other eight criteria for modularity. Knowing its function does not tell us whether its operations are mandatory (1): we might have to endeavour to judge social status, or we might not. The same goes for speed (3); that a mechanism only does one thing does not tell us how fast it does it. That a mechanism's outputs are only of a certain kind says nothing as to whether or not the mechanism itself and/or the mechanism's ontogenetic development are innately specified (5 & 6). That the mechanism seeks exclusively to do one thing seems to have no bearing whatsoever upon whether its internal operations are transparent or opaque to other mechanisms (8), and how many of them. These are criteria (1), (3) through (6), and (8). It seems relatively obvious that they are in no way related to function specificity. None of the properties they concern are influenced by function specificity in any way.

While it may seem that function specificity does influence informational encapsulation (7) and shallow outputs (2), a simple distinction will reveal that it does not. With regard to these two criteria, knowing a mechanism's function could help in determining whether or not that mechanism meets either of them. Knowing a function and knowing instead that a given mechanism simply has a single function, however, are two very different things; the latter gives us no aid in determining anything concerning either (7) or (2). This is an important distinction, as knowing mechanisms to be function specific is just this latter case. Consider: social status would seem to be a property that would require information from all over the brain to determine, so perhaps a mechanism with the function of determining social status would be more likely to be unencapsulated, but the mere fact that any given mechanism is specific to *any given function* does not seem to be enough to make the same sort of inference. Whether or not a function specific mechanism meets either (7) or (2) seems to be entirely contingent. Thus, even if whether a function specific mechanism meets these criteria is *contingent upon that mechanism's particular function*, this does not mean that merely being function specific is enough to guarantee that these criteria are met. Put in terms of (2), knowing the specific function of a mechanism allows us to infer what kind of outputs it might produce (and thus whether or not they would be shallow), but merely knowing that a mechanism is function specific is not enough to infer that its outputs are shallow. You need to know its specific function, not just that it is specific to a single function. Thus, while it may seem that function specificity (9b) has a bearing on whether or not a mechanism meets any other criterion of modularity, in fact it does not, even if certain specific examples of functions might.

All of this together demonstrates that function specificity (9b) truly is not important to what it means to be modular. In short, this means that premise (I) is established: function specificity both is not a necessary, sufficient, or distinct condition of modularity alone and does not imply any criterion that is. Thus, if evolutionary psychologists aim to forward a brand of massive modularity that is an actual modularity hypothesis, they will need to appeal to at least one other criterion that is a necessary, sufficient, or distinct condition of modularity. In order to show that they fail to do so, we will now review the various formulations of massive modularity provided by notable evolutionary psychologists, beginning with Dan Sperber.

Section 2 – Critiquing Sperber

In his paper titled "The modularity of thought and the epidemiology of representations," Sperber provides a theoretical architecture of the mind that he claims to be a variety of massive modularity. Unfortunately, it seems that he relies entirely on function specificity as evidence for the modularity of his thesis. Although Sperber notes that a module is a mechanism "that works pretty much on its own on inputs pertaining to some specific cognitive domain," (Sperber, 1994, 40) the common-sense examples he uses in an attempt to make massive modularity more palatable seem to confuse the input specificity described in the above quote with output or function specificity. In attacking the view that cognitive evolution trends towards demodularization, and thus domain general mechanisms, Sperber provides us with a simple story about creatures he calls 'protorgs'. In this story, protorgs are small scavengers who find themselves commonly in danger of being trampled by large beasts. The approach of such beasts, however, is signaled by a specific noise, N, *and* specific vibrations in the soil, V, and each of these two different percepts is detected by a different perceptual module. If detection of either N or V alone automatically leads to a flight response, then false positives will occur, as the large beasts are only truly approaching when N and V occur simultaneously. Thus, it would evidently be superior to have some sort of mechanism that can take output from the two modules used to detect noises and vibrations, and when its integration processes determine that N and V are occurring simultaneously, then it would trigger the flight response (Sperber, 1994, 43-44).

The problem is not with this story, but instead the moral Sperber draws from it. Despite noting that conceptual integration is taking place in this third mechanism, Sperber nevertheless calls the mechanism "a clear case of a module," even claiming that it is "a domain specific problem solver" (Sperber, 1994, 44). Evidently, this cannot be the case. This conceptual mechanism takes as input two types of data, one from a perceptual module that detects noise and another from a perceptual module that detects vibrations. The only thing that makes this mechanism "a domain specific problem solver" (Sperber, 1994, 44), then, is that the problem the mechanism is solving is specific to a domain: fleeing from large beasts. The mechanism's outputs are of only one kind: triggering the flight response. Its inputs, however, can be of kinds both auditory and vibrational. After further discussion, Sperber even goes so far as to claim that such a mechanism could evolve again to take as input information from additional perceptual faculties (read: additional kinds of data), while still remaining domain specific. So such a mechanism could take as input visual, auditory, tactile, olfactory, and linguistic data in determining its flight response while still remaining domain specific? Evidently, this does not at all describe an input specific mechanism. In calling such a mechanism domain specific, Sperber is obviously missing the point. As long as such a mechanism deals only with fleeing from large beasts, one could call it domain specific with regard to output (function specific), but unfortunately for Sperber, function specificity is not a distinct and necessary condition of modularity. Again, as Sperber himself notes, a module is a mechanism "that works pretty much on its own on inputs pertaining to some specific cognitive domain," (Sperber, 1994, 40). This seems to suggest that he understands that only input specificity is a distinct and necessary condition of modularity, but by conflating function specificity (9b) and input specificity (9a) in his arguments, Sperber's evolutionary account for the development of a massively modular mind ultimately turns out to be the same as the domain general story he is attempting to argue against.

In providing an example of what he believes would be the truly non-modular variety of the above mechanism, Sperber describes it as drawing inferences related to other purposes beyond simply triggering the flight reaction. In what seems to be a provocatively uncharitable illustration of non-modular thought, Sperber notes that were his flight-triggering mechanism to demodularize, it would come to have "interesting thoughts that are filed away for the future" in addition to triggering a flight reaction after inferring danger (Sperber, 1994, 45). The suggestion here is that such a mechanism would be so busy focusing on things apart from survival that natural selection would inevitably weed it out of the population. Unfortunately, while Sperber calls this expansion in function a shift from domain specificity to domain generality, it evidently is only a shift from functional specificity to functional generality. A mechanism could remain output or function specific while becoming input general, as is actually seen in Sperber's own example discussed in the previous paragraph. Function specificity (9b) simply is not a distinct and necessary condition of modularity, however, and thus function specific mechanisms need not be modules. Since, contrary to what Sperber suggests, domain generality does not have to be functional generality, non-modular thought is not restricted to straw-man mechanisms that busy organisms so much with deep thoughts as to hamper their survival.

Furthermore, given that Sperber also admits that central cognition could never be encapsulated (Sperber, 1994, 49), then it would seem that his massive 'modularity' is not truly modularity at all. Massive modularity theorists assert that all (or at least most) of the mechanisms in the mind are modular in character; by focusing on function specificity to the point of dropping both input specificity (9a) and informational encapsulation (7), Sperber is left with, at best, a hodgepodge of criteria that, either alone or combined, are insufficient to entail modularity "to any interesting extent" (Fodor, 1983, 37). As I argued in chapter 1, section 5, informational encapsulation (7) is the hallmark of modularity – its only necessary, sufficient, and distinct condition. Without it, mechanisms simply are not modules. It is likely that Sperber disagrees with me on this issue, but unfortunately he provides little reason why the term 'module,' suggestive as it is of snap-in, necessarily dissociable components, is applicable to his input general mechanisms that are sufficiently interconnected as to be both dependent upon many other mechanisms, and depended upon by them, for proper functionality.

Section 3 – Critiquing Cosmides and Tooby

While Leda Cosmides and John Tooby do not explicitly mention massive modularity in their article "Origins of domain specificity: The evolution of functional organization," they are evidently attempting to vindicate such a thesis. Elsewhere, these authors have described the same evolved cognitive mechanisms they do here, both calling these mechanisms modules in their other publications and claiming that modules realize both peripheral and central cognition (Cosmides & Tooby, 1992, 163 & 165). Unfortunately, they seem so focused on arguing against more empiricist, general-learning accounts of the mind that they overstate the importance of function specificity, casting it as the central idea behind what it means for a mechanism to be domain specific. In beginning their argument as to "why it is implausible and unparsimonious to assume that the human mind is an equipotential, general-purpose machine," Cosmides and Tooby point towards the alarm calls of Vervet monkeys as an illustration of their point (Cosmides & Tooby, 1994, 89). Apparently, these monkeys have evolved three different alarm calls, one for each of the species' primary predators, each of which elicits a specific variety of evasive behavior when uttered. Cosmides and Tooby go on to note that a "general-purpose alarm call (and response system) would be less effective," and take this as evidence that such a behavior would not evolve (Cosmides & Tooby, 1994, 89-90). Since humans would similarly have had such problems to which they would have applied specific solutions, Cosmides and Tooby claim that it only follows that the human mind consists of many "functionally distinct cognitive adaptive specializations" (Cosmides & Tooby, 1994, 90). In short, what they offer here is an evolutionary argument for massive modularity, predicated on the necessity for the proliferation of cognitive mechanisms with specific, adaptive functions.

Since arguments for function specificity (9b) are not arguments for modularity, however, this whole approach falls flat when it comes to massive modularity, especially because it seems that these alarm call mechanisms simply cannot be domain specific in the way modularity requires (viz., input specific). The three predators these monkeys seek to avoid are snakes, leopards, and eagles, and evidently they can be detected by different means, such as through vision or audition. Thus, these alarm call mechanisms must be able to take as input at least these two different information kinds. Furthermore, 'being a leopard' or 'being an eagle' is not the same as 'being blue'. Leopard-ness or eagle-ness must be inferred from a variety of inputs, and requires a number of inputs from different informational domains (shape, colour, and, naïve biology, to name a few). That being the case, these alarm call mechanisms, while indeed function specific (9b), involve the same kind of conceptual integration that we saw in Sperber's story about protorgs, meaning that these mechanisms must be input general in the same way.

Unfortunately, Cosmides and Tooby do not seem to offer any reason to believe that these adaptive specializations are also input specific (or otherwise interestingly modular). Every time they mention domain specificity, they seem to mean only function specificity; Cosmides and Tooby even seem to define a domain as something like a fitness problem with a unique solution (Cosmides & Tooby, 1994, 92 & 96). If domains are simply unique adaptive problems, then it would seem to follow that domain specific mechanisms are unique cognitive solutions to those problems. Defining domains in terms of function (and not input) precludes the ability to define domain specificity in terms of input. Considering that Cosmides and Tooby also call domain specific mechanisms evolutionary adaptations, their assertion that "natural selection theory is a theory of *function*" (original emphasis) seems to be the final nail in the coffin when it comes to their lack of commitment to input specificity (Cosmides & Tooby, 1994, 96). At most, it would seem that the mechanisms proposed by Cosmides and Tooby are merely output specific (9b), meaning that those mechanisms must meet at least one other criterion that is either a necessary, sufficient, or distinct condition of modularity in order for them to actually be modular.

Taking into consideration the list of example 'modules' provided by Cosmides and Tooby in their article "The psychological foundations of culture," it seems unlikely that such mechanisms meet any of the other criteria of modularity. In arguing that an evolutionary picture of the mind can bring both power and breadth to bear in problem solving, Cosmides and Tooby propose an enormous list of function specific, cognitive adaptations. This list includes “a social-exchange module, an emotion-perception module, a kin-oriented motivation module, […], a child-care module, a social-inference module, a sexual-attraction module, a semantic-inference module, a friendship module,” and many, many others (Tooby & Cosmides, 1992, 113). Many of these would have to be extremely complex mechanisms, and there is very little reason to believe that several of them could meet most of Fodor’s modularity criteria, let alone those that are necessary, sufficient, or distinct conditions of modularity. Take the social-inference ‘module’ as an example; such a mechanism could never be mandatory (1), as the sheer number of inferences that might be made in social situations would inevitably lead to a combinatorial explosion were they mandatorily run. Social inferences could also very easily yield outputs that are anything but shallow (2). If to be a shallow output is to be neither unobservable nor theory-laden (viz., an output that can be inferred from its corresponding input without the aid of any information not contained in what is transduced), then it would seem that most social inferences could never be shallow; power-relations, status, love, friendship, and rivalry are all social phenomena that seem eminently unobservable. That is why we would need a social-*inference* mechanism to point them out.

Such a mechanism could perhaps be innately specified, or have an innately specified ontogenetic development, but these two criteria seem to be the only criteria that a social-inference mechanism could ever satisfy. Even if our theoretical social-inference mechanism was completely innately specified (5 & 6), however, that says nothing as to whether it is modular. As we concluded in chapter 1, section 3, innateness and innately specified development are neither necessary, sufficient, nor distinct conditions of modularity; they are largely orthogonal to the concept. Speed (3) would likely be impossible: social situations are too full of alternative explanations for individual phenomena for inferences to be quickly and easily derived from them. Furthermore, the idea that social inferences are opaque to the remainder of cognition (8) is obviously false. We regularly reason about social relationships consciously, determining who is friendly with who and for what reasons. To say that the architecture of the mechanisms that accomplished these tasks made their internal representations opaque to us would be to say that we do not have access to the flow of the very logic we are reasoning with. Evidently, this is not the case. Finally, social-inference could never be an informationally encapsulated (7) process. That such a mechanism would need to query other mechanisms all over the mind is obvious. There are simply too many disparate data points such a mechanism would need to integrate for it to either function entirely using proprietary transducers/databases, be functionally dissociable (4), or specific to a single kind of input (9a). It seems to me, then, that this mechanism (and many of the other mechanisms Cosmides and Tooby propose to be modules) is anything but a module. Returning to my original point, the mechanisms proposed by Cosmides and Tooby may be function specific, but they simply do not meet any criteria that are necessary, sufficient, or distinct conditions of modularity. Since function specificity is not a necessary, sufficient, or distinct condition of modularity, it therefore seems to be the case that the cognitive adaptations Cosmides and Tooby propose just are not modules.

Section 4 – Critiquing Pinker

In his book *How the Mind Works*, Steven Pinker makes several claims in favour of a variety of massively modular architecture. He asserts that "the mind is organized into modules or mental organs, each with a specialized design that makes it an expert in one arena of interaction with the world" (Pinker, 1997, 21). Evidently, this statement is an expression of adherence to some variety of massive modularity. Given that Pinker also claims that modules were "shaped by natural selection to solve the problems of the hunting and gathering life led by our ancestors in most of our evolutionary history," it would appear that he ascribes to a variety of massive modularity similar to that espoused by the other evolutionary psychologists we have discussed above (Pinker, 1997, 21). In providing support for these claims, Pinker provides an important argument for the thesis, one which some thinkers consider to be quite strong (Samuels, 2000, 31). Unfortunately, Pinker offers a relatively cursory formulation of this argument over a great many pages (Pinker, 1997, 21-30), so I will be providing a formal reconstruction. The argument is another argument from evolutionary theory, and goes as follows:

I. The human mind is a successful product of natural selection.
II. In order to be a successful product of natural selection, the human mind must have evolved to solve a great array of complex adaptive problems.
III. A human mind that consisted primarily of domain general mechanisms would be incapable of solving such an array of adaptive problems.
IV. If the human mind evolved to solve such an array of adaptive problems, then it could not have evolved so as to consist primarily of domain general mechanisms.

C. The human mind must have evolved so as to consist primarily of domain specific mechanisms (read: so as to be massively modular).

Evidently, premise (III) is the key premise, and it has not been accepted by anti-modularists without resistance. Most commonly, it is claimed that a mind consisting primarily of domain general mechanisms, but that came equipped with innate, domain specific bodies of knowledge would be no less effective at solving adaptive problems than the massively modular minds proposed by evolutionary psychologists (Samuels, 2000, 31). The idea here is that the mind at large is equipotential and domain general, but not content-independent, meaning that it possesses plenty of innate knowledge, but is otherwise structured entirely out of mechanisms that can tackle any and every problem, situation, input, or computation with equal capacity. An illustration used to argue for such an approach is made using the human capacity of 'mindreading,' or the ability to attribute mental states to others and explain/predict behavior through appeals to mental states. While some have posited the existence of a mindreading module (Baron-Cohen, 1994, 513-514), others have claimed that there exists a literal Theory of Mind, a database of mental state knowledge that can be accessed (consciously or not) by the mind in order to perform mindreading functions (Gopnik & Wellman, 1994, 257).

According to Samuels, if we assume that even domain specific mechanisms require domain specific bodies of knowledge in order to successfully perform their functions, then positing domain specific mechanisms may be multiplying entities without necessity (Samules, 2000, 33). Put another way, if even domain specific mechanisms require databases of domain specific knowledge, then perhaps it is actually the *databases* that are doing the heavy lifting, and are therefore all that is required in order to be capable of solving adaptive problems. As Pinker himself notes,

> Modern digital computers can do a lot with a little, but that 'little' still includes distinct, hard-wired vocabularies for text, graphics, logic, and several kinds of numbers. When the computers are programmed into artificial intelligence reasoning systems, they have to be innately endowed with an understanding of the basic categories of the world… That is no less true of the human mind, (Pinker, 1997, 316).

As an argument from analogy, what this is essentially claiming is that computers require a great deal of innate knowledge in order to be as effective as they are today. Unfortunately, this argument offers no reason to believe the stronger claim Pinker is trying to make: that not only domain specific bodies of knowledge, but also domain specific mechanisms are necessary in order to successfully solve adaptive problems. Thus, as thinkers like Samuels claim, perhaps we should consider the possibility that, when it comes to an evolved picture of the human mind, domain specific knowledge may be the secret ingredient natural selection has endowed us with, not domain specific cognitive mechanisms (Samuels, 2000, 32-33).

While demonstrating that Pinker's argument justifies only domain specific knowledge is a fine strategy when it comes to resisting massive modularity, there are other reasons to believe that Samuels is wrong in thinking that evolution has built our brains to be domain general computers. Thankfully, the domain general picture of the mind and the massively modular picture are not the only options available to us; there is a middle ground that can both concede Samuels' above point and not fall prey to other arguments against the domain general position. Since those other arguments will appear next chapter, however, I will reserve explaining either them or this middle ground position for later.

Foreshadowing aside, Samuels' approach is not the anti-modularist's only means of resistance. Pinker's entire argument can actually be granted by an opponent of massive modularity without conceding defeat and changing position on the topic. In short, this is because a mind that consists primarily of domain specific mechanisms, if those mechanisms are function specific (9b), need not be modular in construction. Even if one grants (III), the thrust of Pinker's argument – that it is theoretically impossible for a purely domain general human psychology to have evolved - one need not immediately accept massive modularity. In truth, this is because the conclusion of the argument is merely that the mind must consist of domain specific mechanisms. In order to establish massive modularity, there must be an additional premise: something to the effect of 'if a mind consists primarily of domain specific mechanisms, then that mind is massively modular in construction'. The issue here is that such a premise can be interpreted in terms of function specificity (9b) as opposed to the intended input specificity (9a), leaving us with the following: 'if a mind consists primarily of function specific mechanisms, then that mind is massively modular in construction'. Unfortunately for Pinker, when this necessary additional premise is interpreted in terms of function specificity (9b), it is patently false. As I have spent the bulk of this chapter demonstrating, arguments for output specificity (9b) do not amount to arguments for massive modularity.

Section 5 – Concluding remarks

Before we close, I would like to clarify my opinion of evolutionary psychology understood as a family of theories. All of my arguments here have aimed at demonstrating that the 'massive modularity' of evolutionary psychologists is not really a modularity hypothesis at all. This is not to say, however, that I believe the theories themselves to be wrong. My only disagreement with Sperber, Cosmides, Tooby, and Pinker concerns whether or not the mechanisms they propose are modules (which is to say, whether or they are informationally encapsulated, inaccessible, fast, mandatory, functionally dissociable, etc). Thus, as I have said previously, the argument I have provided here is an attempt to disentangle evolutionary psychology and massive modularity, opening the door for an alternative evolutionary account of the mind. As such, I would feel perfectly confident in granting Cosmides and Tooby's various arguments *contra* the traditional general-learning faculty conception of the mind, and I feel that doing so in no way threatens my own criticisms of their position.

In conclusion, after examining the arguments of all these prominent evolutionary psychologists, it seems that the theories they provide, while likely containing valuable insights into the architecture of the mind, are far from modular. As we saw at the beginning of this chapter, function specificity (9b) is not central to modularity. This is because function specificity is neither a necessary, sufficient, nor distinct condition of modularity. All of Sperber, Cosmides, Tooby, and Pinker focus so heavily on arguing for function specificity (9b), however, that they fail to demonstrate that the mechanisms they propose can meet any other criterion that is a necessary, sufficient, or distinct condition of modularity. By focusing on evolutionary arguments, and thus arguments from function, evolutionary psychologists seem to have conflated output specificity (9b) and input specificity (9a), which has led them to consider the former to be central to modularity. Thus, by predicating the modularity of their mechanisms exclusively upon output specificity, evolutionary psychologists have ensured that their brand of 'massive modularity' is not modular.

## Chapter 3 – A critique of Carruthers

Having completed our treatment of the brand of massive modularity adhered to by evolutionary psychologists, we will now move to Carruthers' more sophisticated articulation of the thesis. Unlike the evolutionary psychologists, Carruthers does not rely on functional specificity (9b) in an attempt to establish the modularity of the mechanisms he proposes. Instead, he offers several theoretical arguments in favour of a massively modular architecture, and then describes a specific kind of massively modular architecture that he believes to be plausible given these arguments. Carruthers focuses on Fodor's criteria, just as I have, but drops many of them, as he believes that in order for an entirely modular mind to be plausible, more mechanisms need to qualify as modules. I believe that this approach fails; Carruthers weakens his conception of modularity too much, leaving him with a view of the mind that, by the definition of modularity I established in chapter 1, is decidedly non-modular. In substantiating this critique of Carruthers, I will begin with a treatment of the various arguments he offers in favour of massive modularity, where I will show that those arguments are insufficiently strong to establish even a foothold for massive modularity. Then, I will move on to Carruthers' particular account of massive modularity. By returning to an important argument that I made in chapter 1, section 4, concerning the logical implication from input specificity (9a) to informational encapsulation (7), I hope to ultimately demonstrate that the mechanisms Carruthers' proposes cannot, by definition, be modular. Finally, after completing my treatment of Carruthers, I will make some portentous remarks as to where we are left, assuming that my criticisms go through.

Section 1 – The argument from design

In his book *The Architecture of the Mind*, the first argument Carruthers offers in favour of a massively modular architecture is his argument from design. This argument concerns not only cognitive modularity, but complex functional systems in general; the idea being that the mind most likely has a massively modular organization simply because all sufficiently complex functional systems do. In providing evidence for this claim, Carruthers turns to biology, making the classic analogy between physiological organs and cognitive mechanisms. Essentially, his claim is that since "there is a very great deal of evidence from across many levels in biology to the effect that complex functional systems are built up out of assemblies of sub-components," then such an organization is likely to describe "cognition also, provided that it is appropriate to think of cognitive systems as biological ones, which have been subject to natural selection," (Carruthers, 2006, 13). Thus, he presents the argument from design in the following form (Carruthers, 2006, 25):

I. Biological systems are designed systems, constructed incrementally.
II. Such systems, when complex, need to have massively modular organization.
III. The human mind is a biological system, and is complex.

C. So the human mind will be massively modular in its organization.

Evidently, premise (II) is the key premise, and Carruthers seeks to substantiate it through an appeal to Herbert Simon's two watchmakers, Tempus and Hora. Tempus assembles each watch he makes as an individual unit; the simplest components of the watch are added together one at a time until the entire watch is complete. The issue here is supposed to be that, until the watch is complete, its assembly cannot be interrupted without it falling to pieces. In order to avoid this, Hora assembles each watch she makes out of previously-assembled sets of sub-components; the simplest components are combined first into subsystems, and eventually those subsystems are combined in order to create the finished product (Carruthers, 2006, 13). As Carruthers points out, biological systems are built the same way Hora builds watches. Genes, cells, organs, and organisms are all (at least physiologically) constructed in this incremental manner. In fact, it is this particular variety of organization that makes evolution such a powerful force. With natural selection capable of acting upon any item within any level of the functional hierarchy of a particular capacity, it can fine-tune that capacity to meet a variety of selective pressures without otherwise greatly changing or compromising the system as a whole. At the very least, this incremental organization is a highly common evolutionary solution to the need for specialization (Carruthers, 2006, 14). Whether or not incremental organization is the best solution to the need for specialization is not here established, but the fact that both evolution and human ingenuity have hit upon it as a 'go-to' method suggests that it might be.

What is important, as Carruthers notes, is that this incremental organization would entail, at the very least, that the mind will consist “of systems that are to some significant degree dissociable, and each of which has a distinctive function, or set of functions, to perform” (Carruthers, 2006, 17). He cites evidence suggesting that the mind is rife with dissociable mechanisms, pointing out existing cognitive disorders that impair *only* the capacities to reason about mental states, or to recognize faces, or to name natural kinds (Carruthers, 2006, 18). Furthermore, incremental organization would suggest that various functions of the whole are performed entirely by subsystems. Since the mind fulfills so many functions, this means that there must be a veritable cornucopia of subsystems at work within it, seemingly a form of massive modularity (Carruthers, 2006, 20).

This argument does indeed seem to show that the mind likely consists of many mechanisms, each specific to its own distinctive function(s), which is essentially a form of function or output specificity (9b). As I argued at length in chapter 2, section 1, however, arguments for functional specificity (9b) do not amount to arguments for modularity. As such, the important aspect of this argument is that it also presents the mind as a collection of mechanisms that are *dissociable*. As I concluded in chapter 1, section 5, functional dissociabililty (4) is a necessary condition of modularity, though it is not a sufficient condition, as mechanisms can be functionally dissociable without being independent in the way the term 'modularity' suggests. Thus, the important question here is what variety of dissociability is established by Carruthers' argument from design. If, at its strongest, the argument cannot establish genuine (or necessary) functional dissociability, then it does not prove that his mechanisms are modular.

Unfortunately for Carruthers, his argument from design cannot establish functional dissociability at all, even effective (or accidental) functional dissociability. This is because the thrust of the argument is something like the following: complex systems are inevitably structured as hierarchies of subsystems because only hierarchically-structured systems can remain stable *during the incremental construction of complexity*. As already mentioned, the problem with Tempus' method of watchmaking is that by not constructing his complete system out of subsystems, that system is likely to fall apart along the way if something goes awry. If he were to construct his complete system out of subsystems, however, that would not mean that once it was assembled, those subsystems would not depend upon one another in order to work properly. That is because this argument concerns the incremental *construction* of systems, or structural dissociability, which is something entirely apart from the overall *functioning* of incrementally constructed systems, or functional dissociability (4). It may very well be that the mind is designed such that it can develop incrementally, but it does not follow from this that the mechanisms which make it up are capable of functioning at full capacity before being assembled.

As even Carruthers himself notes, "biological systems like hearts and lungs are closely interconnected with many others, of course – each is tightly tied into the bilateral organization of the body, and presupposes the existence of the other," (Carruthers, 2006, 14). Essentially, Carruthers is admitting here that while hearts and lungs are constructed incrementally, they cannot function without one another; thus, they are structurally dissociable, but not functionally dissociable. Given that the argument from design is an analogical argument from biological architecture to cognitive architecture, we can assume likewise for cognitive mechanisms: that they are incrementally constructed, but can only function at full capacity once that incremental construction is complete. Put another way: the individually-assembled mechanisms in Hora's watches may not fall apart before she finally puts them together, but this does not mean that they are already able to tick. Structural dissociability is not a necessary, sufficient, or distinctive condition of modularity; in fact, it obviously bears so little relation to modularity that we were able to discount it almost immediately when we discussed it in chapter 1. Thus, since Carruthers' argument from design can establish only structural dissociability, when it needs to establish necessary functional dissociability, it does not gain even an inch of ground for massive modularity.

What Carruthers' argument from design does do, however, is provide powerful resistance to more traditional domain general architectures, even those equipped with domain specific bodies of knowledge. In chapter 2, section 4, I discussed Samuels' reply to Pinker's argument for massive modularity. Essentially, this reply pointed out that all that is required in order to be capable of solving adaptive problems are innate bodies of domain specific *knowledge*, meaning that domain specific *mechanisms* are superfluous, and therefore that positing their existence amounts to multiplying entities without necessity. I also noted, however, that the domain general position Samuels adopts in replying to Pinker is vulnerable to another argument that I had not yet mentioned. Carruthers' argument for design is that argument. Consider: in order to be the sort of system that can be constructed incrementally, the mind cannot consist of only a few general-purpose mechanisms that employ domain specific bodies on knowledge in order to solve adaptive problems, as general-purpose computers are sufficiently complex that they need to be built from scratch. Domain general computers are like Tempus's watches: they cannot hold together until complete. Thus, the mind cannot be primarily domain general, as otherwise it could not have been developed incrementally by evolution. As Carruthers noted when he summed up his argument from design, "the prediction of [the argument], then, is that cognition will be structured out of systems, … each of which has a distinctive function, or set of functions, to perform" (Carruthers, 2006, 17).

Thus, while the argument from design may not be an argument for modularity, it is an argument in favour of a mind that consists of a *great many mechanisms*, all of which are function specific (9b). The argument from design demonstrates that, insofar as domain general mechanisms cannot be constructed incrementally, a primarily domain general mind could not have evolved. Thus, the argument from design provides theoretical concerns in favour of a massively function specific (9b) mind, which is evidently not the picture of the mind Samuels is in favour of. As to why all of the above is so important, this is an explanation we will have to get to later, in my final chapter.

## Section 2 – The argument from computational tractability

Returning to the task at hand, Carruthers' other argument in favour of massive modularity is based upon computational tractability. Finding its foundations in the computational theory of mind, this argument holds that, if the mind is a computational system, it follows that the mind must be bound by the same rules as one, and thus that the computations made by the mind must be tractable: both theoretically possible and practically possible (viz., possible given what we know of the human brain in particular). The thrust of Carruthers' argument from computational tractability, then, is essentially that only a massively modular mind could actually perform the functions we know the human brain to be capable of and still be computationally tractable. The argument as he presents it, however, is far too ambitious. Carruthers anticipates this, and weakens it through introducing the concept of wide-scope encapsulation. Including the necessary alterations of several premises to better fit his weakened version, Carruthers' argument from computational tractability is as follows (Carruthers, 2006, 44-45 & 59):

I. The mind is computationally realized.
II. If the mind is computationally realized, then all mental processes must be computationally tractable.
III. Only processes that are at least wide-scope encapsulated are so tractable.
IV. If all mental processes must be computationally tractable and only processes that are at least wide-scope encapsulated are so tractable, then all mental processes must be at least wide-scope encapsulated.
V. If all mental processes are at least wide-scope encapsulated, then the mind is massively modular.

C. The mind is massively modular.

While Carruthers offers some arguments in favour of (I), I see no need to include them here. Instead, we will simply assume the truth of the computational theory of mind for the sake of argument. Carruthers' second premise is that all computational processes must be sufficiently tractable. This is to say, these processes must be such that they can be accomplished by the computational mechanisms of the brain both in theory and in reality; in short, not only must such processes be achievable within a finite period of time, but the human brain must be able to perform them within a timeframe characteristic of actual human cognition. In order to meet this requirement, cognitive mechanisms need to be *frugal* both in the amount of information they use and the complexity of calculations they employ in order to accomplish their functions (Carruthers, 2006, 52). This assertion seems trivially true: if the brain is a computational system, it of course must be capable of effecting whatever computations it is known to, using only its own limited set of resources. The key premises then, are obviously (III) and (V).

Central to both these premises, of course, is the concept of wide-scope encapsulation. Carruthers contrasts wide-scope encapsulation with traditional informational encapsulation (7), which he renames narrow-scope encapsulation. If mechanism M is narrowly encapsulated, then its internal operations cannot make use of a set of information outside of either its input or its own dedicated databank, where 'a set of information' means information used by a set of other mechanisms (including sets of one). If mechanism M is widely encapsulated, however, it must be incapable (or not in need) of querying most of the information in the brain by virtue of being restricted by its own operations in how much information it can query. In short, a narrowly encapsulated system has unrestricted access to a restricted set of information, while a widely encapsulated system has restricted access to an unrestricted set of information (Carruthers, 2006, 58). In both cases, this ultimately means that M, in performing its function, will make use of less information than there exists in the mind at large. Thus, Carruthers' new wide-scope encapsulation appears to be frugal concerning the amount of information it can employ in the same way traditional (narrow) encapsulation is, which is what buys it tractability.

Carruthers appeals to heuristics as a possible means of realizing wide-scope encapsulation – heuristics being simple rules that both guide and limit computations. As an example, he cites a tested heuristic that is simply called 'Take the Best'. This particular heuristic, when presented with a query, will search for the variety of information that, in the past, has been found to provide the most accurate determination. For example, if 'Take the Best' were given a query concerning which of two cities was larger, it might first check which city has the most professional sports teams, and if that fails, which city is seen more often on maps. Ultimately, the heuristic will either find an indicator and employ that, or reach a random determination after meeting the predetermined limit of indicators it may query (Carruthers, 2006, 55). Either way, the heuristic will only ever query a very limited amount of data in any single computation. Thus, it indeed does seem to be the case that heuristically guided mechanisms are widely encapsulated, and thus tractable.

Insofar as searching the contents of the entire mind would indeed be intractable, it seems to me that Carruthers' very weak restriction on the amount of searching that may actually occur provides a fine minimal baseline for establishing tractability. All of this may not decisively demonstrate that all mechanisms in the mind must employ *at least* widely encapsulated processes in order to be computationally tractable, but I do not feel that Carruthers needs to be pressed on this point. My own issue is not with (III), but with (V); in short, I do not believe that widely encapsulated mechanisms are even remotely modular. As we have just seen, wide-scope encapsulation essentially boils down to computational frugality, and not much else. This is because wide-scope 'encapsulation' is not really a form of informational encapsulation at all. A mechanism that is widely 'encapsulated' can still query all the information in the entirety of the mind, just not in a single act of problem-solving. As such, there does not exist a set of information which the internal operations of a widely 'encapsulated' mechanism cannot make use of. Furthermore, consider the implication from informational encapsulation (7) to functional dissociability (4) that I mentioned in chapter 1, section 2. If mechanism M does not make use of any information stored in/outputted by its neighboring mechanism X, then it will suffer *absolutely no* failure in functionality when X does. If M *both can and does* make use of some information stored in/outputted by X, X's failure would leave M without all the information it needs to complete its own operations, leading to a failure in M's functionality as well. Wide-scope 'encapsulated' mechanisms both can and often do make use of information stored in/outputted by any and every other mechanism in the mind, meaning that they are dependent upon those other mechanisms in a way traditionally encapsulated mechanisms are not. Thus, while traditional informational encapsulation (7) implies functional dissociability (4), wide-scope ‘encapsulation’ does not.

It seems relatively clear, then, that wide-scope ‘encapsulation’ is not really a form of informational encapsulation at all. Furthermore, wide-scope 'encapsulation' is neither a necessary, sufficient, nor distinct condition of modularity. We can imagine widely ‘encapsulated’ mechanisms that meet no other criterion of modularity other than speed (3), and this is quite simply because to be widely ‘encapsulated’ simply is to be speedy. As we discussed in chapter 1, section 5, independence or autonomy of functionality is at the heart of the concept of modularity. We could imagine a mind where each and every one of its mechanisms employed some sort of computationally frugal heuristic, but where all these mechanisms were also so massively interconnected that if a single mechanism was damaged, then every other mechanism would suffer from a catastrophic performance loss. Fast mechanisms could deliver complex outputs, have their workings be completely transparent to the remainder of cognition, and accept input of any kind, all while remaining fast and frugal. As such, widely ‘encapsulated’ mechanisms need not have shallow outputs (2), be inaccessible (8), or be input specific (9a). Since computational frugality implies nothing aside from speed, and wide-scope ‘encapsulation’ is merely computational frugality, wide-scope ‘encapsulation’ also implies nothing aside from speed. Unfortunately, while cognitive modules are generally fast (as actual informational encapsulation is also a means of achieving computational frugality), cognitive mechanisms that are fast are not necessarily modules. As such, arguments for wide-scope encapsulation do not amount to arguments for modularity. The take-home moral from all this: premise (V) in Carruthers’ argument from computational tractability, that ‘if all mental processes are at least wide-scope encapsulated, then the mind is massively modular,’ is unsound, and thus the argument from computational tractability is too.

Furthermore, the fact that Carruthers’ wide-scope ‘encapsulation’ is not a necessary, sufficient, or distinctive condition of modularity becomes an even larger problem for Carruthers when he elects to drop traditional informational encapsulation (7) from his picture of what it means to be modular (Carruthers, 2006, 12). Unlike Fodor, Carruthers believes that such capacities as consciousness, logical inference, and practical reasoning (viz. central cognitive processes) are accomplished by modules in the same way as vision, language, or audition. Since Carruthers also believes that such central systems are obviously not encapsulated in the traditional sense, he instead elects to claim that they are wide-scope 'encapsulated'. Unfortunately, as we just saw, Carruthers' wide-scope 'encapsulation' is not at all related to what it means to be modular in even remotely the same way traditional encapsulation is. As a result, Carruthers' commitment to maintaining the input specificity (9a) of his ostensible modules is an issue for his account at large. By eschewing traditional informational encapsulation in favor of his own wide-scope frugality, yet at the same time attempting to keep input specificity, Carruthers' ultimately renders his account either inconsistent or not substantially modular.

Section 3 – Why Carruthers' account cannot be modular

Early on in *The Architecture of the Mind*, Carruthers addresses encapsulation (7), inaccessibility (8), and domain specificity (9). In so doing, he points out (and rightly so) that each of these three criteria, despite how often they are blurred together, are quite distinct. Focusing on input specificity (9a) without making explicit the distinction between it and output specificity (9b) that I explained in chapter 1, he notes that "one can easily envisage systems that might *lack* domain specificity, for example (being capable of receiving any sort of content as input), but whose internal operations are nevertheless encapsulated and inaccessible," (Carruthers, 2006, 5; original emphasis). This is all well and good, and it indeed shows that neither informational encapsulation (7) nor inaccessibility (8) imply input specificity (9a). Unfortunately, this says nothing about the implication *from input specificity to encapsulation* (from 9a to 7). Assuming the argument I made for this implication in chapter 1 is correct, then denying informational encapsulation (7) necessarily means denying input specificity (9a) as well, via *modus tollens*, meaning that Carruthers' attempt to eschew informational encapsulation while remaining committed to input specificity makes his account self-contradictory.

Of course, there are options available to Carruthers that would allow him to avoid this fatal conclusion. That being said, it seems best to at least briefly summarize my argument for the implication from input specificity (9a) to informational encapsulation (7), before moving on to explaining those options – self-contradiction is a pretty sensational accusation, after all. Briefly, then, at its most basic, my argument came in three steps. In the first step, I noted that since being restricted in access to a set of information and being unable to use that information were effectively the same thing, any mechanism that could not use information either stored in or outputted by another mechanism was *effectively encapsulated* from that other mechanism. This is even the case if there exists no access restriction between the two mechanisms in question, meaning that one mechanism can be *effectively encapsulated* from another mechanism despite not being *genuinely* *encapsulated* from that other mechanism. In the second step of my argument, I presented a pair of theoretical cases in order to illustrate the idea I had developed in step 1, and they showed that being unable to process information stored in/outputted by another mechanism, due to input specificity (9a), resulted in the same inability to make use of information as actual informational encapsulation (7). In the last step, I broke down the distinction between information taken as input and information used in processing that input, but only in terms of a given mechanism's informational domain. In short, I argued that a given mechanism must be able to process all the kinds of information involved in transforming input into output, and thus that, since being specific to an informational domain concerns what kinds of information a given mechanism can process, excluding information used exclusively in processing from the mechanism's informational domain is arbitrary and thus unwarranted.

If this entire argument is correct, then being input specific (9a) renders a mechanism incapable of using any information outside of the domain it is specific to, which means that the mechanism will act exactly as if it had been encapsulated (7) from all that information its input specificity makes it unable to compute over. Thus, if this entire argument is correct, then input specificity (9a) indeed implies informational encapsulation (7), which means that if Carruthers is to drop informational encapsulation from his definition of modularity, he must drop input specificity as well, on pain of logical contradiction. Unfortunately, he does not. After completing his own treatment of Fodor's nine criteria of modularity, Carruthers concludes that "if a thesis of massive mental modularity is to be remotely plausible, then by 'module' we cannot mean 'Fodor-module'," (Carruthers, 2006, 12). As such, he strikes several criteria off his own list, but he elects to keep mandatory operations (1), fixed *qua* innate neural architecture (5), inaccessibility (8), *and input specificity* (9a). Assuming that he wants to avoid self-contradiction, then, one option available to Carruthers is to alter his account so that it denies input specificity (9a) as well. This would ultimately leave Carruthers with only criteria (1), (5), and (8) remaining. Unfortuantely, this is a very big problem for Carruthers, simply because these remaining criteria, either individually or collectively, are insufficiently important to what it means to be modular to establish a non-trivial thesis of modularity.

Before moving on to Carruthers' next option, I will explain why the above is the case. Consider: it is perfectly coherent to conceive of a mechanism that is mandatory (1) but yet exhibits none of the other criteria of modularity. Such a mechanism would engage of its own accord, but then be capable of querying any and all information in the mind, slowly producing complex outputs, and in the process making its operations available to any other mechanism that might attempt to access them. Insofar as this mechanism would be so strongly interconnected to the rest of the brain, it obviously could not be modular. As I argued in chapter 1, section 5, and have returned to many times, massively interdependent interconnectivity is simply not what to be modular is; it is practically its opposite.

Fixed neural architecture seen as innateness (5) is surely possible without modularity. Chomsky argued for prolific innateness without positing modular mechanisms while easily avoiding internal contradiction (Fodor, 1983, 3-4). Universal inaccessibility (8), while it would indeed be a strong and interesting claim for any theory to make, is nevertheless also not enough to turn a theory into a modularity hypothesis. As with mandatory operation or innate neural architecture, inaccessibility could very easily be a characteristic of a cognitive mechanism that is otherwise completely non-modular. Put another way, a non-modularity theorist could easily agree with any and all of these criteria without having to change his or her stance on modularity.

This is because, either alone or combined, these criteria simply do not entail the level of independence suggested by the term 'module'. A module is, by definition, a piece of cognitive machinery independent of cognition at large; unlike informational encapsulation (7), none of these remaining criteria imply the independent or autonomous functioning that the term 'module' requires. When he himself introduces what he calls a *weak* species of modularity, Carruthers points to a simple hi-fi system, noting that it is modular insofar as various pieces can be removed without effecting the operations of their neighbors (Carruthers, 2006, 2). Even all of (1), (5), and (8) together do not entail mechanisms that meet that description.

Ultimately, this means that in giving up every criteria but mandatory operations (1), innateness (5), and inaccessibility (8), Carruthers' thesis stops being a thesis of massive modularity – it stops being a thesis that non-modularists need disagree with.  No encapsulation (7) means no input specificity (9a), and denying these central criteria (especially alongside so many others) effectively means denying modularity.  I am inclined to think that this is why Carruthers appeals to wide-scope encapsulation, but insofar as this 'encapsulation' is actually only a variety of computational frugality (Carruthers, 2006, 58-59), it also does not entail the independence of functionality that the term 'module' requires.  In short, modularity requires a specific variety of necessary autonomy that can only be achieved by traditional, narrow-scope informational encapsulation (7), and none of (1), (5), or (8), individually or together, can guarantee that, even with wide-scope 'encapsulation' at their side.  As Dustin Stokes and Vincent Bergeron claim, in their article "Modular architectures and informational encapsulation: A dilemma," modularity without encapsulation "is no different from what all cognitive scientists take for granted" (Stokes and Bergeron, 2012, 29).  Evidently, this is not a path Carruthers would want to pursue.

Unfortunately, the only other means of avoiding self-contradiction available to Carruthers is equally unattractive.  Instead of denying input specificity (9a) as the result of denying informational encapsulation (7), Carruthers could elect to commit to the encapsulation of his mechanisms, keeping both (7) and (9a), which would make his mechanisms unquestionably modular.  As he already recognizes, however, this would also make those mechanisms unquestionably implausible. Carruthers notes that "a [universally encapsulated] architecture isn't even plausible as an account of the minds of insects and arachnids, let alone of the human mind," (Carruthers, 2006, 7).  This is likely true for a great many reasons, but I will not discuss them now.  Instead, I will offer an extended argument against the plausibility of a universally (or at least primarily) encapsulated conception of the mind next chapter.  For now, that Carruthers himself refuses to accept a primarily encapsulated mind as feasible is enough to demonstrate that this means of avoiding self-contradiction is also one he is likely uninterested in pursuing. This looks like checkmate for Carruthers.

Section 4 – Conclusion

Before I close this chapter it seems best to both provide a recap of our critique of massive modularity thus far, and briefly discuss where we stand in light of that critique. My criticisms of both evolutionary psychology and Carruthers were of a similar vein. After arguing for a definition of modularity in chaper 1, I then turned to Sperber, Cosmides, Tooby, Pinker, and Carruthers, demonstrating that, according to that definition, each and every one of them ultimately proposed a variety of massive 'modularity' that in fact was not modular at all. While the evolutionary psychologists leaned too heavily on functional specificity (9b) in hopes that it would buy them the modularity of their mechanisms, Carruthers watered down his definition of modularity too much, losing himself every necessary, sufficient, and distinctive condition of modularity in the process. This became particularly obvious given the implication I established in chapter 1 from input specific (9a) to informational encapsulation (7). Finally, over the course of this whole critique, I addressed several theoretical arguments for massive modularity. In so doing, I demonstrated that, while those arguments did indeed provide strong reasons to dobut a domain general picture of cognition, they were insufficiently strong to establish a thesis as radical as massive modularity.

In the end, it has become clear that both the arguments for massive modularity and the various formulations of the thesis we have discussed all suffer from the same endemic problem: the picture of the mind that they propose is not actually a *modular* one. Whether it is because he has recognized this endemic issue, or merely because he refuses to stretch his position beyond what he can justify, Pinker presents a relatively modest view of the mind by the end of *How the Mind Works*, arguable changing his tune and denying his adherence to the massive modularity hypothesis. As he notes,

> The word 'module' brings to mind detachable, snap-in components, and that is misleading… Mental modules need not be tightly sealed off from one another, communicating only through a few narrow pipelines… Modules are defined by the special things they do with the information available to them, not necessarily by the kinds of information they have available. So the metaphor of the mental module is a bit clumsy; a better one is Noam Chomsky's 'mental organ', (Pinker, 30-31).

In his response to Fodor's critique of *How the Mind Works*, Pinker makes it clear that his loyalties lie with a more modest picture of the mind than massive modularity, granting that perhaps some "input systems are domain general with respect to reasoning systems, feeding a number of them with information about objects or bodies or actions" (Pinker, 2005, 17). He acknowledges that this view is one that denies massive modularity, but holds nevertheless that "it does not speak against the mind as a network of subsystems that feed each other in criss-crossing but intelligible ways – the organ system metaphor on which [*How the Mind Works*] is based" (Pinker, 2005, 17). Thus, despite the fact that he offered what appeared to be an argument for massive modularity, it seems that, in the end, Pinker might not be a massive modularist at all. Instead, his position seems focused on the importance of evolutionary concerns and massive functional specificity (9b). As he writes, mental organs "accomplish specialized functions, thanks to their specialized structures, but don't necessarily come in encapsulated packages" (Pinker, 1997, 315).

Pinker's more modest approach is particularly important in light of my own position. As I briefly mentioned near the end of my treatment of evolutionary psychology, I believe that, in between massive modularity and more traditional domain general approaches to cognition, there is a middle ground that can both benefit from all the various arguments for massive modularity I have discussed thus far *and* not fall prey to arguments targeting domain general positions, such as Carruthers' argument from design. Furthermore, I believe that this middle-ground position can do all of this without even flirting with a massively modular conception of the mind. This is the fruit born of disentangling massive modularity from the various insights that belong to the theories of Pinker, Carruthers, Sperber, Cosmides, and Tooby: a massively function specific (9b), but non-modular view of the mind that can take advantage of all the various arguments discussed so far without requiring them to be any stronger than I have shown them to be. Given Pinker's responses to Fodor's criticisms of *How the Mind Works*, this may very well be what many of these thinkers have been arguing for all along. Even if so, however, their formulation of the idea has been sufficiently ambiguous and *confused with* massive modularity that I still consider disentanglement of the two positions to have been a valuable endeavor. Thus, after putting the final nails in the coffin of massive modularity in the next chapter, I will spend the final chapter of this project explaining and arguing in favour of this massively function specific, middle-ground position.

## Chapter 4 – Why the mind cannot be universally encapsulated

In my last chapter, after finishing my discussion of Carruthers, I promised an extended argument in favour of why neither he nor any other massive modularist could plausibly claim that the mind is composed primarily of informationally encapsulated mechanisms. This argument is particularly important for the following reason: according to the definition of modularity I established in chapter 1, a mechanism *must* be informationally encapsulated (7) in order to be modular. Therefore, a successful argument to the effect that ‘informational encapsulation cannot be a virtually universal property of the mind’ amounts to a decisive demonstration that *there is no possible formulation of massive modularity that is both truly modular and plausible*. This is a decisive argument against massive modularity, and not just against any individual author’s particular brand of massive modularity, but of massive modularity as a hypothesis at large. Substantiating this argument is the goal of this chapter, and in doing so, I will target Zenon Pylyshyn's cognitive impenetrability of perception thesis, since to be cognitively impenetrable is to be informationally encapsulated (either genuinely or effectively). The way I see it, the best case for cognitive impenetrability lies in early vision; thus, if early vision turns out to be penetrable, it is likely inevitable that the rest of cognition is as well. Therefore, since the denial of cognitive impenetrability is the denial of informational encapsulation, a strong case for the cognitive *penetrability* of early vision essentially amounts to a strong case against virtually universal informational encapsulation. Ultimately, I believe that such a case can be made, and thus that massive modularity is a doomed enterprise that can never hope to live up to its name and be truly modular.

Section 1 – The cognitive impenetrability thesis

In his article "Is vision continuous with cognition? The case for cognitive impenetrability of visual perception," Zenon Pylyshyn describes a perceptual mechanism as cognitively impenetrable if it "is prohibited from accessing relevant expectations, knowledge, and utilities in determining the function it computes" (Pylyshyn, 1999, Abstract). For a mechanism to be informationally encapsulated (7), that mechanism must be unable to make use of a specific set of information outside of either its input or its own dedicated databank (if it has one). Both input specificity (9a) and genuine access restrictions are means of achieving this state of affairs; thus what Pylyshyn is advocating in describing cognitive impenetrability is a means for individual cognitive mechanisms to be informationally encapsulated from higher order cognition (expectations, knowledge, utilities). In specifically claiming that early vision is cognitively impenetrable, he is denying the causal efficacy of top-down effects upon such processes as identifying motion, size, and colour. In short, Pylyshyn is asserting that certain top-down effects cannot change what early visual processes actually see (Pylyshyn, 1999, 343-344).

In addition to defining cognitive penetrability, Pylyshyn's article seeks to precisely circumscribe the kinds of top-down effects that qualify as genuine cases of the cognitive penetration of early vision. While not explicitly spelled out, Pylyshyn's two restrictions for cognitive penetration in the case of early vision are as follows: in order for a top-down effect on early vision to qualify as a genuine case of cognitive penetration, it may not originate from a cause that occurs prior to early vision (A), nor may it originate from a cause that occurs after early vision (B) (Pylyshyn, 1999, 343-344). Restriction (A) seeks to rule out the influence of attention as a case of cognitive penetration. Attention most definitely changes what we see, but by directing what we look at, not by altering what individual percepts looks like. In cases of change blindness, for example, subjects still perceive the whole scene before them; they simply do not notice certain aspects of it. One person noticing different aspects of a single visual scene than another does not cause that first person to actually see a different visual scene, holds Pylyshyn. Thus, cognitive penetration of early vision is not occurring here.

Restriction (B) seeks to rule out decisions made or pattern recognition accomplished based on what we see. Pylyshyn provides an example of recognizing an acquaintance: realizing that a single individual we can see is someone we know does not actually alter what we are seeing, despite the fact that it does alter our visual experience of the scene before us. In recognizing Ms. Jones in a crowd, we go from experiencing a crowd of strangers before us to experiencing Ms. Jones amidst a crowd of strangers. This recognition indeed appears to be a top-down effect, insofar as judging a specific person to be Ms. Jones could "depend on anything [we] know about Ms. Jones and her habits as well as her whereabouts and lots of other things" (Pylyshyn, 1999, 344). This judgment does not qualify as cognitive penetration of the early visual system, however, because it does not change our percept of the visual scene before us. We still see the same thing, even if we think of it in different terms.

By proposing these two restrictions, Pylyshyn equips his position with several examples that seem to militate in favour of the cognitive impenetrability of early vision. The most well-known of all his examples is the Müller-Lyer illusion, which, not incidentally, is also Fodor's prime example of and most distinctive piece of evidence for informational encapsulation (Fodor, 1990, 249-250). As both Pylyshyn and Fodor point out, knowing the truth about the Müller-Lyer illusion (viz. that the lines are in fact of equal length) does nothing to make the illusion disappear. The idea here is that while individual beliefs can change depending on your other beliefs (the reliability of your sources, your important goals, your motivation to be skeptical, and other such considerations), what you see in the case of the Müller-Lyer illusion, and perceptual illusions in general, does not change depending on your beliefs, "even when what you know is both relevant to what you are looking at and at variance with how you see it" (Pylyshyn, 1999, 344). For both Pylyshyn and Fodor, this is taken as evidence of an important difference between perception and cognition – evidence of an informational access restriction between early vision and central cognition. The Müller-Lyer illusion, then, seems to be evidence for the cognitive impenetrability (or informational encapsulation) of early vision.

Despite how strong Pylyshyn's case may initially appear, many authors have brought together evidence and arguments that cast it into doubt at essentially every level. In what follows, I will provide a brief review of several articles that argue in favour of the cognitive penetrability of early vision, beginning with some psychological experiments that seem to provide examples of top-down effects upon perception that are not ruled out by Pylyshyn's restrictions. Afterwards, I will turn to a paper by Mark Rowlands that, based on an extended view of perceptual processes, argues for denying Pylyshyn's two restrictions (A & B) in such a way that allows additional top-down effects to qualify as genuine cases of cognitive penetration. Finally, I will turn to a psycho-anthropological study that claims to demonstrate that the Müller-Lyer illusion is relative to culture, and thus perhaps also relative to learned, unconscious visual assumptions (read: a kind of knowledge that would seem to be exactly the variety Pylyshyn denies can affect early visual processes).

Section 2 – Counterexamples to the cognitive impenetrability thesis

In their paper, "Modular architectures and information encapsulation: A dilemma," Stokes and Bergeron cite experimental evidence from New Look psychologists Jerome Bruner and C.C. Goodman. In the experiment, subjects turned dials in order to adjust the size of light beams shining upon circular objects in front of them. They were told to match the size of the circle of light with the size of the object the light was shining upon. When those objects were coins of currency as opposed to worthless discs, subjects adjusted their light beams to be on average larger, despite the coins and discs in fact being identical in size; subjects in lower income brackets did this almost universally (Stokes and Bergeron, 2012, 12-13). As beliefs concerning value or worth are paradigmatic illustrations of information provided by higher cognitive processes, this experiment seems to decisively demonstrate that the early visual process of size-perception is cognitively penetrated by central cognition (Stokes and Bergeron, 2012, 12-13). Furthermore, matching the size of the light beam to the size of the coin or disc indeed seems to be a process that violates neither of Pylyshyn's restrictions. It involves a stationary viewpoint, as the beam of light was shining directly onto the coin itself; thus, shifts in attention cannot be the cause of differing percepts. Furthermore, since the subjects needed only adjust the size of the light beam to match it to the coins/discs, all in real time, right in front of them, neither memory nor post-perceptual decision-making/recognition could have come into play. Being a certain size within one's field of view is evidently not the same sort of thing as being Ms. Jones. Therefore, one must conclude from this study either that subjects were systematically misjudging the size of coins, but not discs, seemingly for no reason, or that subjects were actually seeing the coins as physically larger. The first interpretation is obviously *ad hoc*; it postulates a baseless systematic misjudgement, seemingly for the sole purpose of rejecting the second interpretation. As such, it seems to be that this is a genuine case of the cognitive penetration of early visual (early peripheral) processes: belief concerning monetary value is actually changing how subjects perceive size.

In another article entitled "Cognitive penetration of colour experience: Rethinking the issue in light of an indirect mechanism," Fiona MacPherson cites another experiment, this one from psychologists J.L. Delk and S. Fillenbaum. In this experiment, subjects adjusted the colour of a background along the spectrum between yellow and red so as to match the colour of shapes placed atop the background. When those shapes were suggestive of stereotypically red things, such as apples or hearts, subjects adjusted the background more towards the red side than they did when the shapes were simple, geometric circles or squares, despite the fact that all the shapes were actually of identical colour (MacPherson, 2012, 38-39). Again, since this process involved no attention shift, no recall, and no decision-making or pattern recognition, it either implies that subjects were systematically misjudging the colour of some shapes, but not others, for no reason, or that subjects were actually seeing the shapes as more red in cases where those shapes suggested stereotypically red things (MacPherson, 2012, 38-39). Once again, this seems to be a pretty clear-cut example of the cognitive penetration of early visual process: beliefs or expectations concerning characteristic colour is changing how subjects perceive actual colour.

Stokes and Bergeron assert that "a perceptual experience E is cognitively penetrated if and only if (1) E is causally dependent upon some cognitive state C and (2) the causal link between E and C is internal and mental" (Stokes & Bergeron, 2012, 10-11). If one makes the following two assumptions, this definition offers a succinct description of what Pylyshyn calls genuine cognitive penetration. First, 'mental' must mean something informational, as otherwise neurochemicals would meet all the criteria, and I doubt Stokes and Bergeron intended that. Second, it must be the case that attention is not a cognitive state. If so, changes in E dependent upon attention are ruled out here, as Pylyshyn would want them to be. As for recognition, one need not even make these assumptions for it to be ruled out; recognition does not actually alter E, meaning that E is not causally dependent upon recognition. Finally, this definition rules out trivial cases of where the cognitive state of 'wanting to go see a particular film' causes us to experience that film as opposed to something else, simply through giving us cause to get up, go to the theatre, and buy a ticket. This is not an entirely internal and mental causal chain, and thus it also does not qualify (Stokes and Bergeron, 2012, 10-11). What is particularly important, however, is that this definition rules out neither of the two experiments mentioned above, just as neither experiment falls prey to Pylyshyn's two restrictions. It seems, then, that these experiments are genuine cases of the cognitive penetration of early vision – solid evidence that even early vision is not informationally encapsulated.

Section 3 – Rowlands' argument for the cognitive penetrability of perception

Even if the conclusions of these experiments could be denied or explained away by a cognitive impenetrability theorist, there are still other reasons to doubt the cognitive impenetrability thesis. In his paper "The cognitive penetrability of perception," Mark Rowlands offers an extensive critique of two assumptions concerning the workings of perception that he believes to be necessary for the coherence of the cognitive impenetrability thesis. As Rowlands asserts, the issue concerning perception is whether, or to what extent "the processes that properly constitute visual perception [are] modified by information input from non-perceptual cognitive vehicles," or more familiarly, to what extent the architecture that realizes visual perception is informationally encapsulated (Rowlands, 2005, Abstract). Rowlands believes that specific properties of perceptual experience suggest that two particular assumptions concerning the workings of perceptual vehicles that underlie the cognitive penetrability thesis may be false. Specifically, Rowlands believes that *the extended character* of perception entails the denial of two assumptions concerning the workings of perception – assumptions which Rowlands claims are the foundations for Pylyshyn's restrictions circumscribing cases of genuine cognitive penetration (Rowlands, 2005, Abstract).

Rowlands calls these two assumptions *the internality assumption* (IA) and *the assumption of genuine duration* (DA). The internality assumption (IA) is essentially just what its name might suggest: it holds that the process of perception "is internal in the sense that it is implemented by, or realized in, structures and mechanisms that do not extend beyond the skin of the perceiving organism" (Rowlands, 2005, 14). Rowlands notes that this would be the case even if the content of perceptual representations is externally determined, as the external individuation of representations' content does not entail the external location of those representations (Rowlands, 2005, 14). The assumption of genuine duration (DA) claims that the processes of perception can be said to start and stop at determinate times, likely beginning and ending coextensively with the process of transformation of the retinal image into a representation (Rowlands, 2005, 15). This assumption also holds, however, that gaps in the activation of perceptual processes always mark new, distinct perceptual tokens. Holding both of these claims in tandem allows us to point to individual examples of perception that genuinely have durations, hence the name (Rowlands, 2005, 15).

Rowlands believes that while it is possible to identify some parts or components of perception that meet these two assumptions, to hold that they constitute the whole of perception is both arbitrary and likely false. He asserts that this follows from the nature of perceptual experience, namely its extended nature (Rowlands, 2005, 19). *Perceptual experience* violates both the internality assumption (IA) and the assumption of genuine duration (DA), he claims, and the way in which it does so means that at least some of the *perceptual processes* which realize it must as well (Rowlands, 2005, 19). If this is the case, then the fact that some perceptual processes violate these assumptions means that perception at large cannot be said to always meet these assumptions (Rowlands, 2005, 15). If perception at large must always meet these assumptions for the cognitive impenetrability thesis to be plausible, then Rowlands' argument from the extended nature of perceptual experience poses a significant threat.

His argument itself is that perceptual experience is extended both spatially and temporally insofar as, instead of internally representing the whole visual scene before us, we represent only pieces of the scene as we look upon them. As he puts it, "we are aware of a complex and detailed world, but this means only that we are aware of the world as being complex and detailed," (Rowlands, 2005, 22-23). In order to experience the specific complexity and detail of any given scene, we need to look out towards what surrounds us, as particular complex and detailed objects are simply not represented internally. Essentially, Rowlands is arguing for the same view of perceptual representation that Daniel Dennett outlines in his book *Consciousness Explained*: that while any action of introspection upon our perceptual experience immediately returns the intuition that 'we have in our heads a complete picture of what we currently see in front us', this is not actually the case.

Dennett clarifies this idea by appealing to a particular eye-tracker experiment, wherein a subject reads words on a computer screen while hooked up to a machine that analyzes his or her eye movements. This machine is so quick and so precise that, in the time it takes a subject to shift his or her focus from one word to the next, it can both determine exactly which word on the screen will become the subject's next focal point and *change that word before the subject can even finish shifting their gaze to focus upon it* (Dennett, 1991, 360-361). For the subject, the experience is perfectly normal; it feels just like reading the text on any other computer screen. Anyone watching the process, however, will see words all over the screen constantly changing and rearranging, going from narrative to nonsense in literally the blink of an eye. Recounting his own experience with this kind of experiment, Dennett writes that "While I waited for the experimenters to turn on the apparatus, I read the text on the screen. I waited, and waited, eager for the trials to begin. I got impatient. 'Why don't you turn it on?' I asked. 'It *is* on,' they replied," (Dennett, 1991, 361; original emphasis). Evidently, this experiment shows that our experience of 'having in our heads a complete picture of everything we currently see in front us' is in fact an illusion. As Dennett so aptly points out, "the representation of presence is not the same as the presence of representation," (Dennett, 1991, 359); we simply do not represent the whole visual scene before us – we only represent whatever piece of the scene that we are focusing on at that very moment.

Rowlands appeals to this 'grand illusion' conception of perception and perceptual experience in order to substantiate his argument in favour of an external view of perceptual processes, turning first towards the phenomenon of change blindness. In these experiments, either a group of local disturbances or a single global disturbance accompanies a change in a presented visual scene. Perhaps a detail changes, just as many blots appear as decoys (the local disturbances variation), or the whole scene flickers briefly as it changes (the global variation). The interesting thing about these experiments is that subjects generally fail to notice that the change has occurred when it is accompanied by such disturbances. While this may be old hat for some, Rowlands draws from it a novel conclusion: if we internally represent the visual scene before us in its full detail and complexity, then change-blindness ought not to occur. If we represented the world in this way, we would simply be able to compare our new percept to our previous one, and that would be enough to notice the differences, at least more often than not (Rowlands, 2005, 21-22). Given that subjects also fail to notice changes they do not attend to in inattentional blindness experiments, it seems that while we experience the world as rich, this is not because we represent what makes it rich, but only its property of richness. As Rowlands ultimately asserts, “external complexity and detail is not internally reproduced” (Rowlands, 2005, 22). Our perceptual experience might suggest otherwise, but given that we seem to be able to instantly attend to any aspect of the visual scene before us in order to scrutinize any detail, it is easy to see how we might come under the impression that we are constantly seeing everything when, in fact, we are not (Rowlands, 2005, 23).

Assuming this is correct, argues Rowlands, perception must be extended in nature. By replacing complex and detailed internal representations of entire visual scenes with series of simpler, more focused representations supplemented by constant visual exploration of the surrounding environment, this approach renders representations themselves insufficient for the production of visual experience (Rowlands, 2005, 24). As such, both internal representations and what Rowlands calls external ‘information-bearing structures’ (which is really just a sophisticated term for external objects of visual perception) are necessary for complete visual experience. Thus, argues Rowlands, since the necessary exploration of such external structures is spatially extended into the environment, so are the processes which constitute visual perception (Rowlands, 2005, 24). By the same token, for Rowlands, the importance of exploring these external information-bearing structures extends perception temporally. As (at least partially) an activity of exploration, perception does not simply begin with a retinal image and end with a transformed percept; instead, perceptual processes are ongoing, continually occurring as long as the organism is actively exploring the environment.

Furthermore, assuming perception is an ongoing activity, temporary periods of inactivity become gaps in the process of perception, not determinant beginnings and endings. Temporarily ceasing exploration need not mark new, distinct perceptual tokens. We may pause during the exploration of a particular scene, but it remains the same scene when we begin exploring it anew (Rowlands, 2005, 24-25). This is because exploration, as an activity, need not take place during a continuous period of time. For example, the activity of sculpting a statue of a person could take place over weeks, months, or even years, and much of that time would evidently be spent engaged in other pursuits. Doubtlessly, during our time spent sleeping, we would not be sculpting the statue, but just as evidently, starting another round at sculpting the statue after a good night's sleep would not amount to starting a whole new sculpture. Pausing during the activity of sculpting need not mark a new, distinct sculpture token. Insofar as Rowlands' view conceives of perception as the activity of exploration, a similar claim can be made regarding it: pausing during the activity of perceptual exploration need not mark a new, distinct perceptual token. Thus, a single ongoing act of perception can have gaps, just as any activity can.

If all the above arguments are true, then this externalized account of perception must obviously lead us to deny both the internality assumption (IA) and the assumption of genuine duration (DA), the assumptions which Rowlands claims to be the foundations underlying Pylyshyn's restrictions circumscribing cognitive penetration. As both spatially and temporally extended, perception would neither be internal (as per IA) nor would it have a genuine, gapless duration (as per DA). As Stokes and Bergeron pointed out in their definition of cognitive penetration, it must be the case that "the causal link between [experience] and [cognition] is internal and mental" (Stokes & Bergeron, 2012, 10-11). If perception is not entirely a spatially internal affair, however, then we cannot reasonably maintain that anything that penetrates it must be entirely internal. Furthermore, Pylyshyn's two restrictions seek to rule out as genuine cases of cognitive penetration effects whose causes occur both before and after perceptual processes; if perceptual processes have no discreet beginning and end, however, and cognition occurring during gaps in perceptual exploration could alter subsequent percepts, then it seems to follow that Pylyshyn's restrictions are false or arbitrary (Rowlands, 2005, 25-27). If perceptual processes are ongoing, with intermittent lacunae, then we simply cannot point to processes that occur before or after perception; other cognitive processes only occur either during perception or amidst it – within gaps of inactivity.

Section 4 – Evaluating Rowlands' argument

A great deal, then, seems to hang on the truth of Rowlands' account. Despite arguing for both the spatial and temporal extension of perception, however, Rowlands admits that establishing only the temporal extension of perception (viz. that perception violates the assumption of genuine duration) is enough to undermine the cognitive impenetrability thesis of Pylyshyn (Rowlands, 2005, 19). As long as we cannot distinguish between processes occurring before/after perception and processes occurring during perception, then Pylyshyn's attempts to rule out attention, pattern recognition, and judgement as causes of genuine cognitive penetration ultimately fail. Thus, even if one can present a powerful argument to deny the spatial extension of perception, the meat of Rowlands' disagreement with Pylyshyn still remains. I say all this at least in part because I do not find myself that convinced by Rowlands' arguments for the spatial extension of perception. It seems to me that in arguing for the spatial extension of perception, Rowlands is guilty both of assuming that the attributes of content can be indicative of the attributes of vehicles and of committing what, in their article "Why the mind is still in the head," Fred Adams and Ken Aizawa call the coupling-constitution fallacy.

Briefly, the thrust of Adams and Aizawa's claim that there is a fallacy is as follows: objects or elements necessary to the proper functioning of a system need not be constitutive of the system. Instead, it could merely be the case that such elements need to be coupled with the system in a certain way for it to function (Adams & Aizawa, 2009, 16-17). Essentially, this is an assertion to the effect that, for example, the mere fact that some source of water is necessary for a sprinkler system to properly douse flames does not mean that the water itself is constitutive of the sprinkler system; the reservoir that holds the water might be, or perhaps the pipes that connect the sprinkler system to the city's water supply, but the water itself is not a proper part of the system. To believe that the water is a constitutive portion of the sprinkler system is to believe that the water is doing productive work, when it is simply what is being worked upon. In his argument for the spatial extension of perception, Rowlands asserts that purpose of perception is “the production of visual experience of the surrounding environment” (Rowlands, 2005, 24). We can think of the perceptual system, then, as a system that produces visual experience. As Adams and Aizawa are (rightfully) pointing out, that Rowlands’ ‘information-bearing structures’ are necessary for that productive process does not make them a part of it. The perceptual system uses them to produce a visual experience in the same way the sprinkler system uses water to douse flames. Therefore, while ‘information-bearing structures’ are indeed necessary for visual experience, such structures are not constitutive of the perceptual system. They must be properly coupled with our internal processes in order for those processes to successfully produce visual experience, but it does not follow from this that these external elements are constitutive in the actual production of visual experience. Therefore, as these external elements are not constitutive of the perceptual system, they have no impact upon whether or not the system at large is spatially extended.

Given this consideration, I find myself very sceptical of Rowlands’ argument for the spatial extension of perception. Fortunately, his argument for the temporal extension of perception takes a different form. Instead of claiming that external elements play a constitutive role in perceptual processes, he simply asserts that since at least a part of those processes takes the form of continual exploration of indeterminate duration, the totality of perception, insofar as it involves temporally extended portions, must be temporally extended in a similar manner. Taking this into consideration (perhaps alongside the fact that temporal extension seems far less radical a conclusion), I see no reason to be sceptical about Rowlands’ argument for the temporal extension of perception.

All of that now said, the takeaway moral from Rowlands’ paper is that temporal extension of perception casts into doubt the two restrictions that Pylyshyn believes circumscribe all genuine cases of cognitive penetration. As an example of how this is the case, consider attention. Pylyshyn’s restriction (A) states that in order for a top-down effect to qualify as a genuine case of cognitive penetration, it may not originate from a cause that occurs prior to early vision. This restriction is considered to apply to cases where attention directs perception. If Rowlands’ arguments for the temporal extension of perception are correct, however, then pinpointing exactly what counts as ‘prior to early vision’ is impossible. As an ongoing, but not continuous process, perception can only be interrupted by cognition, not preceded by it. As such, cases of attentional shift would count as cognitive penetration, and ruling them out for the sake of the cognitive impenetrability thesis would be both question-begging and arbitrary (Rowlands, 2005, 27). Pylyshyn’s restriction (B), that in order for a top-down effect to qualify as a genuine case of cognitive penetration, it may not originate from a cause that occurs after early vision, would be similarly rejected. Ultimately, this means that Rowlands’ argument opens the floodgates wide for many phenomena to qualify as genuine cases of cognitive penetration, including cases as ubiquitous and simple as attentional shifts. Evidently, this would leave the cognitive impenetrability thesis without a leg to stand on.

Section 5 – Evidence for the penetrability of the Müller-Lyer illusion

In case you are not convinced by Rowlands' argument, however, there remains one final paper to which I would like to turn. In their article "Susceptibility to the Müller-Lyer illusion, theory-neutral observation, and the diachronic penetrability of the visual input system," Robert N. McCauley and Joseph Henrich bring together evidence from across seventeen cultures that seems to demonstrate that cultural factors have an influence on perceptual experience, with differences present for the first twenty years of subjects' lives drastically affecting how they perceive the Müller-Lyer stimuli (McCauley & Henrich, 2006, Abstract). I will not summarize their findings here. However, the authors conclude that they "demonstrate considerable cultural variation with respect to humans' susceptibility to the Müller-Lyer illusion and that it looks as though, depending upon their visual experiences during childhood, some people are not susceptible to the illusion at all" (McCauley & Henrich, 2006, 95). Given that the data seem to show that hunter-gatherers from Kalahari Desert are essentially immune to the illusion, I find myself inclined to agree with McCauley and Henrich's interpretation of the results.

The only hitch is that immunity to the illusion is gained over time. As the article's research shows, children are, on average, much more susceptible to the illusion than adults. When certain cultures boast adults immune to the illusion, this seems to be because those adults *learned* their immunity over the first twenty years of their lives (McCauley & Henrich, 2006, 94-95). When it comes to cognitive impenetrability and informational encapsulation, Fodor has only ever denied that higher order cognition can have any effect in the here and now (McCauley & Henrich, 2006, 89). As McCauley and Henrich interpret their results concerning the Müller-Lyer illusion, however, perception of it is *diachronically* penetrated, or cognitively penetrable given enough time. Simply learning that the two lines in the Müller-Lyer illusion are of equivalent length will not cause them to appear that way instantly (which would suggest *synchronic* penetration). At this juncture, the issue for cognitive impenetrability theorists becomes whether or not *diachronic* penetration and *synchronic* penetration are two distinct kinds of penetrability. Since the only difference between them seems to be a measure of time, it seems unlikely that the type-distinction between them is nothing more than arbitrary (McCauley & Henrich, 2006, 89). Furthermore, if you do find yourself convinced of Rowlands' temporal extension of perception, then effectively all perception becomes *diachronic* in nature, forcing a denial of the distinction between *diachronic* and *synchronic* penetration at square one.

Even if one manages to draw a solid line between *diachronic* and *synchronic* penetrability, however, one fact still remains: the Müller-Lyer illusion is ultimately penetrable, even if only through a long-term process. As McCauley and Henrich stress, what their evidence shows, "even with respect to Fodor and Pylyshyn's favorite example, is that informational encapsulation [or cognitive impenetrability] is not comprehensively specified by the human genome, that it is not pervasive, and that there is no consensus about the pertinent stimuli among human observers" (McCauley & Henrich, 2006, 98). In one way or another, even the Müller-Lyer illusion seems to be subject to cognitive penetration. Since, as stated, both Fodor and Pylyshyn turn to this illusion as their paradigm example of *im*penetrability, these findings give us much reason to doubt the cognitive impenetrability thesis, and thus that informational encapsulation is a prolific property of the mind.

In conclusion, it seems to me that cognitive impenetrability, and thus informational encapsulation, is far from being a universal (or even virtually universal) property in early vision. Courtesy of Stokes, Bergeron, and McPherson, we seem to have before us at least two examples of cognitive penetration that even Pylyshyn himself would call genuine. Furthermore, given that Rowlands gives us reason to doubt the two restrictions Pylyshyn uses to circumscribe cases of genuine cognitive penetration, there seems to be no wiggle room left for the cognitive impenetrability theorist. Thus, to agree with Pylyshyn, one must somehow account for the counterexamples I have given above, which is difficult. By contrast, adopting the picture of cognition that disagrees with Pylyshyn suddenly allows a whole gamut of effects to qualify as genuine cases of cognitive penetration. Finally, regardless of how one deals with this issue, McCauley and Henrich's data stand as a final counterexample to the impenetrability thesis. Restrictions or not, the *diachronic* penetration of Pylyshyn's model example, the Müller-Lyer illusion, remains an enormous problem for the thesis. At the very least, we can say that a primarily informationally encapsulated (7) conception of early vision has been cast into doubt. Given that early vision is informational encapsulation's best bet, it seems very unlikely indeed that the mind at large is primarily encapsulated (or cognitively impenetrable). As even Pylyshyn himself notes, "most psychological processes are cognitively penetrable, which is why behavior is so plastic and why it appears to be so highly stimulus-independent" (Pylyshyn, 1999, 343).

## Chapter 5 – A picture of the mind

Having discussed Cosmides, Tooby, Sperber, Pinker, and Carruthers, and now Pylyshyn's cognitive impenetrability thesis, our critique of massive modularity is complete. Over the course of that critique, however, I repeatedly foreshadowed the importance of several arguments and insights, claiming that they would return for an encore when I proposed my own architectural account of the mind. In this final chapter, I will finally bring resolution to all those portentious remarks by painting an alternative picture of the mind – one that borrows many aspects from the work of the writers I have criticized over the course of this project, but combines them in what I believe to be a powerful way. I will begin this chapter with the central claims of my account, laying them out as a group. I will then address each of these claims, one by one, providing explanations and arguments for each. Fortunately, all of the arguments I provide will already be familiar; since my position in this chapter is a middle-ground position between more traditional domain general accounts and massive modularity, certain arguments both for massive modularity and against domain general cognition can be turned to my benefit. I will not use material only from what we have already covered, however. In the latter portion of this chapter, I will turn to Vincent Bergeron's work *Cognitive architecture and the brain: Beyond domain specific functional specification*, drawing from it several important insights that I believe make my own position both clearer and more realistic. Ultimately, I hope to provide a picture of the mind as an evolved, massive set of function specific mechanisms, that despite being primarily domain specific, is decidedly non-modular in organization.

Before we begin with the main discussion of this final chapter, it seems best to briefly turn back to a particular distinction I made early on that will play a central role in what follows. In chapter 1, section 3, I addressed the criterion of domain specificity. In so doing, I noted that domain specificity could be broken down into two distinct concepts: input specificity (9a) and output/ functional specificity (9b). I later addressed each of these concepts in detail, tackling input specificity in chapter 1, section 4, and functional specificity in chapter 2, section 1. From these discussions, one particular conclusion stood out: while input specificity (9a) implies informational encapsulation (7), and is thus a sufficient condition of modularity, functional specificity is neither a necessary, sufficient, nor distinct condition of modularity alone and does not imply any criterion that is. Thus, function specific mechanisms, despite being domain specific, need not be modular; given that I aim to propose a primarily domain specific, but non-modular picture of the mind in this chapter, this conclusion is of utmost importance.

Section 1 – Foundations of the picture

The central claims supporting this position are as follows: (I) the mind is a product of natural selection; (II) as a product of natural selection, the mind has evolved in such a way as to solve adaptive problems; (III) since evolutionary development is necessarily incremental in nature, the mind must be made of structurally dissociable mechanisms, as otherwise it could not have been constructed by evolutionary processes; (IV) since a mind that consists primarily of a few domain general mechanisms cannot be structurally dissociable, the mind must consist primarily of many domain specific mechanisms instead; and finally, (V) since to be input specific (9a) would make the mind's many mechanisms into modules, and massive modularity simply is not tenable, those mechanisms must instead be function specific (9b). Taken together, these claims do indeed seem to leave us with an evolved, massively domain specific, but non-modular mind. Let us see how well they stand up, then.

I believe that claims (I) through (III) are common ground amongst architectural theorists. As Cosmides and Tooby so aptly put it, "the human brain did not fall out of the sky, an inscrutable artifact of unknown origin," (Cosmides & Tooby, 1994, 85). Much like the heart, the lungs, the eye, or any other complex biological system, the mind must have evolved. Put simply, this is because chance events "do not systematically build intricate and improbably functional arrangements" (Cosmides & Tooby, 1994, 86). Without some organized mechanism to guide it, random chance does not naturally order the chaos of the universe into self-sustaining and self-replicating machines suited so well to surviving and thriving in the environments they find themselves occupying. Thus, insofar as the mind is a complex biological system, it is absolutely obvious that it must be the product of evolutionary development. So much for claim (I), that the mind is a product of natural selection.

Of course, if the mind is a product of evolution, then it must have developed in such a way as to solve adaptive problems. Simply put, this is because 'to evolve' and 'to develop in such a way as to solve adaptive problems' are virtually one and the same thing. The theory of natural selection posits that only those organisms properly adapted to the environment they find themselves in will survive. Thus, that humanity is alive today is proof that the human species has developed the proper adaptations in order to become fit, meaning that everything about humans, more or less, has developed in such a way as to solve adaptive problems. In short, as a *successful* product of natural selection, it necessarily follows that the human mind is (largely) adaptive. Thus is established the truth of claim (II), that as a product of natural selection, the mind has evolved in such a way as to solve adaptive problems.

Claim (III), that 'since evolutionary development is necessarily incremental in nature, the mind must be made of structurally dissociable mechanisms, as otherwise it could not have been constructed by evolutionary processes', is relatively complex, and thus requires some unpacking. Essentially, the idea behind this claim concerns the process of evolutionary development itself. Consider: evolution is not a foresightful process; organisms do not evolve in anticipation of a threat upon their survival, but in reaction to it. As a result, evolution builds organisms bit by bit, adding adaptation on top of adaptation, each in reaction to the latest adaptive problem. Evolution cannot build complex functional systems from scratch, because "natural selection had no crystal ball," (Cosmides & Tooby, 1994, 87). In short, because evolutionary development is reactionary and historical, as opposed to foresightful, all of its products (the mind included) are not designed as elegant machines capable of solving as many different problems as possible with the fewest resources. Instead, evolved systems are just big jumbles of adaptations - coalitions of individually constructed mechanisms, each a solution specified to a particular problem.

Importantly, this means that evolution must construct complex functional systems incrementally, or to put the same idea in different words: the systems constructed by evolutionary processes must be structurally dissociable. If we think of this in terms of Simon's watchmakers, Tempus and Hora, evolution must build organisms the same way Hora builds watches. Hora assembles each watch she makes out of previously-assembled sets of sub-components; she first combines the simplest components into subsystems, and eventually she combines those subsystems in order to create the finished product. She does this because if she were to assemble each watch she makes as an individual unit, like Tempus, adding the simplest components of the watch together one at a time until the entire watch is complete, then any interruption would cause the *incomplete* watch to fall to pieces, as only the finished design is intended to hold together. To complete the analogy, then, evolution must build as Hora does, because *evolved systems are always incomplete*; there is no final, structurally sound design evolution is building towards. An organism is a constantly adapting system, which means that new mechanisms are constantly being added to that system. Therefore, the organism needs to be constructed in such a way that, despite being an eternal work-in-progress, it will not die (fall apart) for lack of certain components. Evolved systems must be structurally dissociable.

This above argument is essentially Carruthers' argument from design, or at least what that argument became after our critique. Despite being posited by Carruthers as an argument in favor of massive modularity, our discussion of the argument from design demonstrated that it was actually an argument against more traditional, domain general accounts of cognition, like that of Samuels'. As I pointed out during that discussion, incremental organization implies both domain specificity and dissociability, but it does not necessarily imply input specificity (9a) and functional dissociability (4). The argument from design only warrants structural dissociability and functional specificity (9b), which is important because while functional dissociability (4) and input specificity (9a) are necessary and sufficient conditions of modularity, respectively, structural dissociability and functional specificity (9b) are neither necessary nor sufficient.

However, being uninteresting in terms of modularity does not make structural dissociability and functional specificity uninteresting in general. As I appropriated Carruthers' argument from design in order to demonstrate, because it is an evolved (and thus structurally dissociable) system, the mind cannot consist of only a few domain general mechanisms that employ domain specific bodies of knowledge in order to solve adaptive problems. This is because such mechanisms are exactly the sort of systems that must be built from scratch as elegant machines capable of solving as many different problems as possible with the fewest resources, and as the argument from design tells us, evolution cannot construct such systems. As such, it follows that cognition must consist of mechanisms that each have "a distinctive function, or set of functions, to perform" (Carruthers, 2006, 17). Since the mind fulfills so many functions, this means that there must be a veritable cornucopia of mechanisms at work within it. All of this leaves us, therefore, with an argument to the effect that the mind must both be structurally dissociable and massively function specific (9b). Thus, Carruthers' argument from design can be taken as evidence in favor of both claim (III), that 'since evolutionary development is necessarily incremental in nature, the mind must be made of structurally dissociable mechanisms, as otherwise it could not have been constructed by evolutionary processes', and claim (IV), that 'since a mind that consists primarily of a few domain general mechanisms cannot be structurally dissociable, the mind must consist primarily of many domain specific mechanisms instead'.

That leaves us with claim (V), that since to be input specific (9a) would make the mind's many mechanisms modular, and massive modularity simply is not tenable, the many domain specific mechanisms of the mind must instead be function specific (9b). Evidently, this claim is supported by every preceding chapter of this project. Specifically important to this claim, however, is my argument that functional specificity (or output specificity, 9b) is neither a necessary, sufficient, nor distinct condition of modularity, which appeared in chapter 2, section 1. This argument is important because if you accept it, then you also accept at least the logic of claim (V), that if the mind must be primarily domain specific, but cannot be massively modular, then it must be primarily function specific (9b). Whether you grant that the mind must be primarily domain specific turns on whether you grant everything I have argued above in this chapter, and whether you grant that the mind cannot be massively modular turns on whether you grant everything I have argued otherwise. Obviously, then, claim (V) is the lynchpin of my entire architectural account. As such, I will spend the remainder of this chapter focused on explaining why this massively function specific (9b) picture of the mind is both beneficial to adopt and likely not as radical as it may at first sound, all in hopes of providing additional reason to agree with claim (V).

Section 2 – Dissecting functional specificity

To provide support for claim (V), I now turn to Vincent Bergeron's piece, *Cognitive architecture and the brain: Beyond domain specific functional specification*. In this article, Bergeron breaks down the idea of functional specificity in two ways, drawing one distinction between two understandings of function and another distinction between two understandings of specificity. He then uses these distinctions to establish the possibility of domain specific, but non-modular mechanisms, much like Michael Anderson or Tim Shallice (Anderson, 2010; Shallice, 1988). When tackling cognitive functions, he separates them into *cognitive workings* and *cognitive roles*. A mechanism's working, he states, is the cognitive operation (or set of operations) the mechanism *performs*. To perform an operation, according to Bergeron, is to carry it out in full, meaning that a working is some very specific process that takes only a single mechanism to fully complete (Bergeron, 2008, 65-66). To turn back to a classic example, consider a clock. It is the function of the clock to 'tell time,' but telling time is not the function of any single mechanism in the clock. Individual components might have functions such as 'move the second hand forward by one notch', 'count the movements of second hand in groups of sixty, sending signal S as each new group is completed', and 'upon receiving signal S, move the minute hand forward by one notch'. Such simple operations are the *workings* of various mechanisms in the clock.

The *role* of these mechanisms, however, is simply to tell time. A cognitive mechanism's role is the cognitive operation (or set of operations) the mechanism *participates in*. To participate in an operation is not to carry it out in full, but instead to be part of a group of mechanisms that together realize that operation, carrying it out in full as a complete coalition. As Bergeron points out, we are aware of several areas in the brain that play a part in facial recognition. To say that a given mechanism plays such a part is to say that facial recognition is at least one of that mechanism's cognitive roles (Bergeron, 2008, 66-67). He also points out, as I have gestured toward with my clock example, that this distinction between two different conceptions of function is not specific to cognitive mechanisms alone. The rhythmic contractions of the heart are an operation it carries out fully as a single organ, and are thus its working, whereas circulating blood through the body is an operation the heart is integral to, but does not complete on its own, and is thus its role.

When tackling specificity, Bergeron draws a distinction between mechanisms that are domain specific and mechanisms that are operation-specific. Domain specific mechanisms can perform many different kinds of operations, but only using a single kind of information, whereas operation specific mechanisms can only perform a single kind of operation, but using many different kinds of information (Bergeron, 2008, 67-69). As an example, two mechanisms involved in memory, M and X, would be domain specific if M could only compute over object memory while X could only compute over facial memory. If M and X could each compute over both object and facial memory, but M only stored information in memory while X only retrieved information from storage, then both mechanisms would be operation-specific.

Having made these distinctions, Bergeron uses them to establish two types of domain specific, but non-modular cognitive mechanisms – exactly the sorts of mechanisms I claim the mind might primarily be made of. These mechanisms are both presented as relationships between form and function: the form-working relationship and the cognitive working zone. At its simplest, a cognitive working zone is a region (or perhaps several regions) of the brain that realize a single cognitive mechanism, one which fully carries out some cognitive operation(s). That this mechanism is capable of fully carrying out its operation(s) is a direct result of structure, or form, be that physical or virtual (Bergeron, 2008, 69-70). This means that any given cognitive working zone's form and function are related to one another in the same way regardless of other variables. Contrast this to the relationship between a mechanism's form and its cognitive role, or the function it participates in. A mechanism's form cannot determine its role in the same way it can determine its working. For example, Broca's area could be involved in realizing several cognitive capacities, such as 'understanding sentences', 'counting', and 'melody recognition'. Each of these capacities would therefore be one of the various cognitive roles of Broca's area. This means, however, that the function of Broca's area, when understood in terms of its cognitive role, would change depending on cognitive context. However, if Broca's area performs only a single cognitive operation (parsing linearly-arranged strings of information), then its working remains invariable regardless of context. This is because its working is what it does, while its role is what it is involved in; the former is determined by its structure, and does not change unless the mechanism's structure does as well, whereas the latter is determined by its context, which only very loosely depends upon the mechanism itself (Bergeron, 2008, 69-70).

The form-working relationship is essentially the same as the cognitive working zone, except that it narrows down the operation(s) of the mechanism, specifying them in terms of informational domain(s), instead of describing them generally (Bergeron, 2008, 70). For example, instead of claiming that Broca's area parses linearly arranged strings of information, we might say that it 'performs linguistic parsing', 'performs musical parsing', and 'performs numerical parsing', thus assigning it several functions, each of them corresponding to a different informational kind. The structures that Bergeron calls form-working relationships are still regions of the brain which perform specific workings as a direct result of their architecture; the only difference is that their workings are described in terms of traditional informational domains (such as language, music, or numbers). The important thing is, even if a given mechanism performs functions only within a single informational domain, that does not *necessarily* make it domain specific *qua* input specific (9a), and therefore does not necessarily make it modular. If Broca's area is capable of performing only linguistic parsing, that does not necessarily mean that it can use only linguistic data in performing that function; it may make use of both linguistic information and expectation information, but only output linguistic information. That would make it domain specific, but only domain specific *qua* output specific (9b). The same is true if the working of Broca's area is instead thought of as parsing linearly-arranged strings of information; it may only parse certain kinds of information (kinds that can be linearized), but it could make use of many additional kinds of information over the course of doing so. Thus, specifying form-working relationships or cognitive-working zones does not amount to specifying modules.

Section 3 – Benefits of our dissection

The upshot of all the above is as follows: by breaking down functional specificity and providing two non-modular, but domain specific (function specific) types of mechanism, Bergeron has established a massively function specific (9b) conception of the mind that has a great degree of plausibility. As an illustration, I ask that you once again consider the clock: in order for the whole system to tell time, many subsystems must perform subservient functions - many of which, when taken out of context, may appear to be orthogonal to the main function of telling time. The same is true of any complex function performed by the mind. Take, as an example, incest avoidance. Evidently, avoiding incest is an important step towards creating fit offspring that can carry on your genetic code. It is also, however, a very complicated step; in order to avoid incest, an organism needs to be able to recognize conspecifics, determine which conspecifics it is related to, and have some sort of innate avoidance reaction that kicks in when a possible dalliance is determined to be incestuous. Furthermore, these individual pieces themselves are relatively complex. Determining relatedness alone could be accomplished using many different kinds of evidence; shared upbringing, similarity of features (especially facial features), and even simple linguistic communication could all play a part.

The point of all this is that incest avoidance is sufficiently complex that it could not have evolved in a single step. Just as the eye consists of a lens, rods, cones, a retina, and much else besides, incest avoidance is likely accomplished by a whole coalition of lesser mechanisms. Thus, one *cognitive role* of those lesser mechanisms is incest avoidance. The *cognitive workings* of those mechanisms, however, are whatever laundry list of minor computations is necessary in order to successfully avoid incest, each computation being the working of a single mechanism. Thus, every one of those mechanisms that together realize incest avoidance could be function specific (9b). Consider: if as part of incest avoidance, a single mechanism takes as input a whole variety of data from other mechanisms, and uses it to determine relatedness, then it must be output specific (9b).

While this example is evidently a simplification of the overall process of incest avoidance, it nevertheless serves well enough to prove my point: since its function is to determine relatedness, this hypothetical relatedness-determining mechanism outputs only relatedness information. It outputs only information of a single informational kind (relatedness), meaning that it is output specific, or function specific (9b). Even if the output of this mechanism is coopted to perform many tasks (viz., the mechanism has many cognitive roles), the output is still only of a single informational kind, and the mechanism is still doing only one thing. It is therefore possible for the entire mind to be made of mechanisms that perform only a single working, or a single function, but where those workings are put to different uses, or used in realizing different cognitive roles.

Put another way, conceived of as cognitive working zones or form-working relationships, individual mechanisms become function specific (9b) *because* they are operation-specific. To be correctly described by either of Bergeron's form-function relationships is simply to perform some operation as a direct causal result of having a particular architecture, either physiological or virtual (Bergeron, 2008, 69-70). Most importantly for my purposes, this means that if you agree with Bergeron that the mind likely consists of many cognitive working zones/form-working relationships, then you must hold that the mind is massively function specific (9b). Broca's area is a possible example of a cognitive working zone (Bergeron, 2008, 70-71). Therefore, given that it seems likely that the mind is constructed of many mechanisms of approximately equal complexity to Broca's area, the above argument shows that it is equally likely that the mind is massively function specific (9b). Essentially, in describing his two form-function relationships, Bergeron has provided us not only with a means for mechanisms domain specific while remaining non-modular, but also with a means for all the mechanisms of the mind to be plausibly function specific (9b).

In part, this plausiblity comes from discarding the conceptual baggage from something as unlikely as modules. At the same time, however, Bergeron's argument makes a massively function specific mind appear feasible by demonstrating how few requirements mechanisms must meet in order to be function specific. This second idea is even more obvious if one considers more closely what exactly counts as an informational domain. Once again, consider the example of Broca's area. On the one hand, it could be said to concern three informational kinds: linguistic, musical, and numerical. On the other hand, however, one might claim that since Broca's area outputs only linearly-arranged strings of information, it concerns only a single informational kind. Evidently, the suggestion here is that 'linearly-arranged strings of information' could be considered to be a single informational kind. Put another way, if the cognitive working of Broca's area (roughly) is *only* to parse linearly-arranged strings of information, then it might be considered to be domain specific to the informational domain of 'linearly-arranged strings of information'. As another example, consider a hypothetical mechanism M which performs only the operation of 'detecting edges in the visual field', but is put to use in recognizing faces, identifying inanimate objects, and identifying living things. One could claim that this mechanism concerns three informational domains (faces, objects, and naive biology), but its seems that one could also claim that this mechanism concerns only one informational domain (edges).

This idea makes use of Bergeron's distinction between cognitive workings and cognitive roles in order to demonstrate just how easily even very simple mechanisms can be function specific. While both Broca's area and the above hypothetical mechanism M are concerned with several informational domains via the various cognitive roles each of them plays, both of these mechanisms seem to be specific to only a single informational domain at the level of their workings. Broca's area outputs only linearly-arranged strings of information. Mechanism M outputs only perceived-edge information. Those edges perceived by mechanism M may be later used to identify faces, objects, and living things, but this work is done by other mechanisms – M only detects edges. Thus, if we allow 'edges' or 'linearly-arranged strings of information' to qualify as valid informational domains, then both Broca's area and our hypothetical mechanism M are output or function specific (9b). Each of these two example mechanisms outputs information of only a single informational kind. The informational kinds of 'edges' or 'linearly-arranged strings of information' might be more finely grained than what we think of as traditional informational kinds (such as music or faces), but that is no reason to deny that they are indeed informational kinds.

Consider: vision, audition, pitch, and colour are all used as relatively commonplace examples of informational kinds. The domain of 'colour', however, is a more finely grained domain than that of 'vision', and the domain of 'pitch' is more finely grained than the domain of 'audition'. This is because 'colour' is an informational domain within the greater domain of 'vision', and the same can be said of 'pitch' and 'audition', respectively. Put another way, these more finely grained domains are the building blocks of the informational domains that are more traditionally discussed. Just as cognitive capacities are likely large systems realized by many smaller subsystems, traditional information domains might be coarsely-grained informational kinds composed of many smaller, more finely-grained, 'building block' informational kinds. Morphemes are combined to create words, words are combined to create sentences, and sentences are combined to create paragraphs. Perhaps subdomains are combined to create domains in the same way.

As Cosmides and Tooby point out, natural selection “is not limited to what could be validly deduced by one individual, based on a short period of experience, it is not limited to what is locally perceivable, and it is not confused by spurious local correlations” (Cosmides & Tooby, 1994, 94). It seems fair to assume, then, that evolution has designed our cognitive mechanisms to interact with objects more basic or finely-grained than those we commonly consider. If so, then 'building block' domains would indeed seem to be valid informational kinds, meaning that it would be extremely easy for any mechanism to be function specific (9b), as any mechanism that performs only a single working (is operation specific) would necessarily output information of only one kind. This is the case even if that operation ultimately plays a part in many cognitive roles. As long as ‘building-block’ domains such as edges or linearly-arranged strings of information can just as easily be informational kinds as more traditional domains, such as numbers or colours, then any operation-specific mechanism would necessarily be output specific (9b) to the informational kind it operates upon. To illustrate this idea with a purely fantastical example, if a given mechanism’s function were to 'translate its inputs into binary bits', then while it might output information about numbers, music, or colours, it would still output all of that information in binary bits. Insofar as another mechanism may or may not be able to take binary bits as input, it would seem that they are a distinct informational kind. The same goes for information arranged as linearly-arranged strings of information. Thus it seems that any mechanism in the mind can be at least very weakly function specific, meaning that a massively function specific (9b) mind is easily achieved.

Of course, there are additional reasons beyond palatability to accept this relaxed conception of functional specificity (9b). As Bergeron asserts, one of the greatest advantages of focusing on cognitive working zones is capacity for integration. By attempting to describe the operation a given mechanism fully performs, as opposed to merely the role it plays in a larger cognitive capacity, we can bring together insights and findings from different branches of cognitive research in a new way. Perhaps a hypothetical mechanism M is involved at once in the voice recognition process and the object recognition process. Instead of simply claiming that M performs vocal recognition operation V or object recognition operation O, we could perhaps end up integrating findings to conclude that M is in fact a pattern comparison engine, a small computer that determines how closely observed particulars match generalized representations. This is what Bergeron means when he claims that "the specification of a cognitive working zone will often have the potential to account for the basic contribution of a given brain structure to a wide range of (sometimes seemingly unrelated) cognitive operations" (Bergeron, 2008, 85). By focusing primarily on the elemental contribution of each individual mechanism, prior to the functional roles those mechanisms play, we can also find new ways to relate portions of the mind. The same can be said of what I have somewhat informally called 'building-block' domains. By focusing on the elements that make up more traditional informational domains, we may find that those domains share more in common than we originally thought. If different cognitive capacities turn out to require similar operations in their realization, then knowing which mechanisms perform those operations and what other capacities they are involved in realizing could lead us to find new connections between all those capacities involved.

By the same token, if different cognitive capacities turn out to make use of similar informational elements in their realization, then knowing which mechanisms share what elements in common could lead us to find new connections as well.

This same idea is present when Bergeron notes the boon cognitive working zones are for the evolutionary approach. Finding homologues of human mechanisms in other animals, especially related animals such as primates, gives us much greater insight into the path of human cognitive evolution. Focusing such an investigation on cognitive workings and 'building block' domains is a good way to find similarities between human brains and animal brains. The same cognitive workings can be involved in different cognitive roles, and the same 'building block' domains in different traditional domains. Thus, shifting our focus toward cognitive workings and 'building block' domains amounts to shifting our focus toward the stable elements human and animal brains could share. Put another way, the cognitive role of a given mechanism can change so easily, even without its working changing at all. Thus, comparing mechanisms in terms of their workings, as opposed to their roles, is a far more reliable method for tracing the phylogenetic development of those mechanisms (Bergeron, 2008, 100-102). The same goes for 'building block' domains, as opposed to the more traditional domains.

## Section 4 – Conclusions and further work

In conclusion, I believe that a massively function specific (9b) picture of the mind has several merits. It can accommodate the weakened version of Carruthers' argument from design, but does not fall prey to massive modularity and the arguments against it. Furthermore, as I showed by appealing to Bergeron's work, this picture of the mind is by no means a radical one. It also seems to provide new and fruitful opportunities for both interdisciplinary integration and evolutionary investigation. All in all, this picture of the mind seems to be a successful disentanglement of massive modularity from many of the ideas that made it so enticing; it has many of the theory's strengths without its weaknesses.

Despite everything that I have done over the course of this project, however, I still feel that I have left the 'real work' unfinished. While determining what the mind cannot be is important, at least insofar as it narrows the scope of possibilities, the ultimate goal is to figure out what the mind is. I have made an attempt at meeting that goal here, but it truly is nothing more than an attempt – a fledgling idea that requires much more in the way of work before it can become something interesting and valuable. It seems to me, however, that the massively function specific picture I have proposed deserves further investigation. It is both powerful in what it can account for and plausible in how easily it can be realized. Thus, if I am ever to continue this project, I believe that the next step would be arguing for the truth of this picture, as opposed to merely the possibility that it could be true.

I believe that turning to thinkers who work at the intersection between psychological studies and architectural theorizing would be the best way to take that next step. Authors like Simon Baron-Cohen, Alan Leslie, and Alison Gopnik are all concerned with the mechanical workings of the various cognitive capacities they study. In so doing, they have all organized the results of many experiments in attempts to create platforms of evidence for the various theories each of them adheres to. Measuring the massively function specific view of the mind against those evidential platforms would be an excellent way to determine just how likely it is to be true.

I also believe that it would be beneficial to more carefully explore how well the massively function specific approach accounts for evolutionary concerns. While they have sometimes been given the cold shoulder, I believe that the main thrust behind the approach of evolutionary psychologists is a valuable one. The mind must have evolved, and thus evolutionary constraints are salient when considering the workings of cognitive capacities. Furthermore, the human mind must have evolved from something, so studying the minds of humanity's closest relatives is important in discovering how the mind was built over generations. Knowing the exact phylogenetic history of the mind would be a great help in distinguishing its various functional systems.

There also remains some critical work yet to be done. Over the course of this project, the bulk of the critique I leveled against massive modularity followed from the definition of modularity I established in chapter 1. Given this approach, it seems that my critique would be more complete were I to restate it using a different (and less demanding) approach to definitions, such as exemplars. As definitions go, those created through the specification of necessary and sufficient conditions are sometimes considered to be overly-strict. If my criticisms of massive modularity could continue to hold true even while using a less restrictive approach to definitions, then it would seem that they would be all the better.

Of course, whether all this work will ultimately lead to the rise or downfall of the massively function specific conception of the mind is not something I can currently say with any confidence. What I can say, however, is this: there are many reasons to doubt both more traditional, domain general accounts and massively modular accounts, and yet both pictures of the mind still have their proponents. As a middle-ground position between these two extreme approaches, the massively function specific picture may very well be better than either of them, but we simply will not know without giving it a great deal of our time. This new position deserves careful thought and consideration, even if initial investigations show it to have a few flaws. After all, no theory in science, cognitive or otherwise, is ever perfect.

# Bibliography

Adams, F., & Aizawa, K. (2009). Why the mind is still in the head. In P. Robbins & M. Aydede (Eds.), *The Cambridge Handbook of Situated Cognition*, 78-95. USA: Cambridge University Press.

Anderson, M.L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences*, 33(4), 245-266.

Bargh, J., & Chartrand, T. (1999). The unbearable automaticity of being. *American Psychologist*, 54(7), 462-479.

Baron-Cohen, S. (1994). How to build a baby that can read minds: Cognitive mechanisms in mindreading. *Cahiers de Psychologie Cognitive/ Current Psychology of Cognition*, 13, 513-552.

Bergeron, V. (2008). *Cognitive architecture and the brain: Beyond domain specific functional specification*. Retrieved from https://circle.ubc.ca/handle/2429/2711

Carruthers, P. (2006). *The Architecture of the Mind*. Oxford: Oxford University Press.

Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Science*, 3(3), 115-120.

Cosmides, L., & Tooby, J. (1994). Origins of domain specificity: The evolution of functional organization. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the Mind*, 85-116. Cambridge: Cambridge University Press.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J.H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The Adapted Mind*, 163-228. Oxford: Oxford University Press.

Dennett, D.C. (1991). *Consciousness Explained*. Boston: Little, Brown and Company.

Fodor, J.A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.

Fodor, J.A. (1987). Modules, frames, fridgeons, sleeping dogs, and the music of the spheres. In J. Garfield (Ed.), *Modularity in Knowledge Representation and Natural-Language Understanding*, 26-36. Cambridge, MA: MIT Press.

Fodor, J.A. (1990). *A Theory of Content and other Essays*. Cambridge, MA: MIT Press.

Gopnik, A., & Wellman, H.M. (1994). The theory theory. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the Mind*, 257-293. Cambridge: Cambridge University Press.

Koch, C. (2004). *The Quest for Consciousness: A Neurobological Approach*. Englewood, Colorado: Roberts and Company Press.

MacPherson, F. (2012). Cognitive penetration of colour experience: Rethinking the issue in light of an indirect mechanism. *Philosophy and Phenomenology Research*, 84(1), 24-62.

McCauley, R.N., & Henrich, J. (2006). Susceptibility to the Müller-Lyer illusion, theory-neutral observation, and the diachronic penetrability of the visual input system. *Philosophical Psychology,* 19(1), 79-101.

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

Pinker, S. (2005). So how *does* the mind work?. *Mind & Language*, 20(1), 1-24.

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, 22, 341-423.

Rowlands, M. (2005). The cognitive penetrability of perception. In A. Raftopoulos (Ed.), *Cognitive Penetrability of Perception: Attention, Action, Strategies, and Bottom-Up Constraints*, 13-30. New York: Nova Science Publishers, Inc.

Samuels, R. (2000). Massively modular minds: Evolutionary psychology and cognitive architecture. In P. Carruthers & A. Chamberlain (Eds.), *Evolution and the Human Mind: Modularity, Language, and Meta-Cognition*, 13-46. Cambridge: Cambridge Univeristy Press.

Samuels, R. (2006). Is the human mind massively modular?. In R.J. Stainton (Ed.), *Contemporary Debates in Cognitive Science*, 37-56. Malden: Blackwell Publishing.

Shallice, T. (1988). *From Neuropsychology to Mental Structure*. Cambridge: Cambridge University Press.

Sperber, D. (1994). The modularity of thought and the epidemiology of representations. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the Mind*, 39-67. Cambridge: Cambridge University Press.

Stokes, D., & Bergeron, V. (2012). *Modular Architectures and Informational Encapsulation: A Dilemma*. Retrieved from http://philpapers.org/rec/STOADF

Tooby, J., & Cosmides, L. (1992). The Psychological Foundations of Culture. In J.H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The Adapted Mind*, 19-136. Oxford: Oxford University Press.